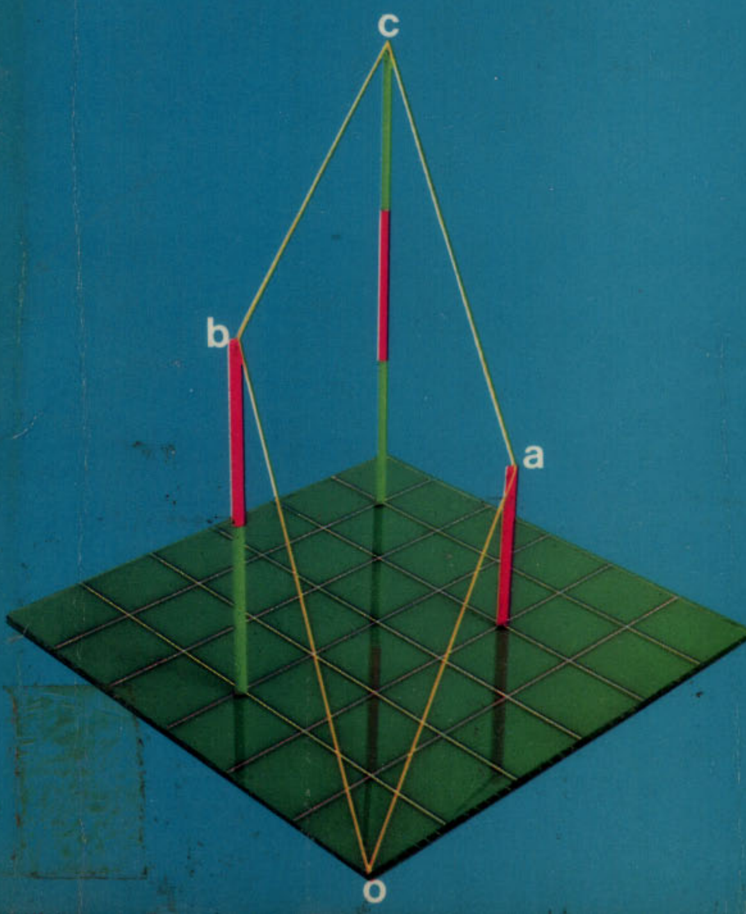
a Pelican Original

Introducing Mathematics: 4

A Path to Modern Mathematics

W.W. Sawyer



W.W. Sawyer A Path to Modern Mathematics

ISBN 0 14
02.0847 X

PELICAN BOOKS

Introducing Mathematics: 4

A PATH TO MODERN MATHEMATICS

W. W. Sawyer was born in 1911. He won scholarships to Highgate and St John's College, Cambridge, where he specialized in quantum theory and relativity. After some ten years spent lecturing in mathematics at various British universities, he went to the College of Technology in Leicester. He became head of the mathematics department, and studied with his colleagues the application of mathematics to industry. In 1948, he became the first head of the mathematics department in what is now the University of Ghana.

For five years from 1951 he was at Canterbury College, New Zealand. He founded a mathematical society for high-school students, and this led to a significant increase in the supply of mathematics teachers in Canterbury province. Professor Sawyer was later invited to help in the reshaping of mathematical education in the U.S.A. He was professor of mathematics at Wesleyan University, Connecticut, from 1958 to 1965. At present he is professor jointly to the College of Education and the Mathematics Department in the University of Toronto.

His books include: *Mathematician's Delight*, *Prelude to Mathematics* (Pelicans), and *A Concrete Approach to Abstract Algebra*. He is co-author of *The Math Workshop for Children* (a textbook for junior American grades), and he has been editor of the *Mathematics Student Journal*. *A Path to Modern Mathematics* is the fourth volume in the Pelican series *Introducing Mathematics* by W. W. Sawyer, of which *Vision in Elementary Mathematics* (Volume One) and *The Search for Pattern* (Volume Three) have already been published. He is married and has a daughter and a grand-daughter.

INTRODUCING MATHEMATICS

4

A Path to Modern Mathematics

W. W. SAWYER

PENGUIN BOOKS

Penguin Books Ltd, Harmondsworth, Middlesex, England
Penguin Books Inc., 7110 Ambassador Road, Baltimore, Maryland 21207, U.S.A.
Penguin Books Australia Ltd, Ringwood, Victoria, Australia

First published 1966
Reprinted 1969, 1971

Made and printed in Great Britain by
Cox & Wyman Ltd,
London, Reading and Fakenham
Set in Monotype Times

Contents

Introduction

IT is highly desirable that the opening pages of a book should give a potential reader some indication of the scope and purpose of the book, its level of difficulty and the knowledge that it presupposes.

First of all, it should be pointed out that, while this book follows *Vision in Elementary Mathematics* in time, it does not follow it in the development of the subject. *Vision in Elementary Mathematics* was aimed at the beginnings of education; it was intended to help the teacher or parent concerned with children between, say, five and thirteen years old; it did not assume any prior knowledge of mathematics apart from that minimum of arithmetic that most people have. This book does assume some background in mathematics. It supposes the reader to be fairly comfortable with the kind of topics covered in my earlier book *Mathematician's Delight*. This does not mean that every chapter requires an understanding of calculus – far from it. If *Vision in Elementary Mathematics* was meant to help the teacher of children five to thirteen years old, this book may be helpful to a teacher who is revising the syllabus for pupils between eleven and eighteen years old. The discussion must make suggestions for the Mathematics and Science Sixth, but it must also concern itself with the eleven-year-olds, and with classes that are not being taught by a mathematics specialist. Chapters One and Three, for example, develop an approach originally published in the *Scientific American*. If this approach is criticized, it will probably be on the grounds that it is too childish. Again, a very considerable part of Chapter Nine has been tried out in schools, and found to be intelligible and entertaining to pupils who knew just a little algebra and Pythagoras' Theorem. Wherever possible, an idea taken from modern mathematics has been explained in terms of quite elementary mathematics.

Should the whole book have been written within an elementary

framework, with all references to calculus excluded? This was decided against for the following reason. I have seen many expositions of modern mathematics which were extremely mystifying. An idea was explained to the audience. The audience were not told where it came from, nor what could be done with it. They had to take it on trust that this was an important mathematical concept, though they could not for the life of them see why. Now mathematics is above all subjects that in which you do not take things on trust; you demand proof. A very poor way to start a campaign for mathematical reform is to brainwash teachers so that they are willing to abandon their critical thinking, and accept changes without knowing why. In no sense can it be said that you are teaching modern mathematics if you simply chip off a few ideas and words from recent mathematics and convey these in isolation, without showing their relationship to other parts of mathematics, the problems they enable you to solve, the reasons why mathematicians attach importance to them.

One would therefore wish to tell a connected story, to show the ideas that led a mathematician to some new concept and the further developments he expected this concept to produce. Now the mathematicians who made the decisive discoveries of the early twentieth century had all had a very thorough training in nineteenth-century mathematics. It was by this that their imaginations had been nourished. Their aims were to clear up those points of logic which the nineteenth century had left obscure, to solve those problems the nineteenth century had left unsolved, to provide neat answers to questions that had been answered clumsily, to penetrate deeply into matters that had been discussed superficially, to unify what had been left separate, to generalize what had been handled as something particular. A twentieth-century discovery would be recognized as significant because of the light it threw on a host of nineteenth-century problems. To present the mathematics of this century without any reference to the previous century is like presenting the third act of a play without any explanation of what is supposed to have happened in the first two acts.

Now mathematics in the seventeenth, eighteenth, and nineteenth centuries is pervaded and dominated by the ideas of the

calculus. One can pick out particular developments – projective geometry, say, and some parts of the theory of numbers – that can be explained without any mention of calculus, but if one were to write an account of mathematics between 1600 and 1900 with all references to calculus forbidden, the work of that epoch would be unrecognizable. There would be unexplained gaps in every chain of cause and effect.

In some countries calculus is not taught in secondary schools at all, or is taught only to a minority, or is taught very late in the syllabus. In such countries there is an almost insoluble problem in presenting modern mathematics in a way that makes sense. Britain is fortunate in that, as the result of prolonged discussions and struggles in the years 1870–1920, calculus is now taught to a very considerable part of our population. This includes not only pupils in the 'academic' streams at secondary schools but engineering apprentices in the National Certificate courses as well. It may be urged that the calculus taught to our sixteen-year-olds does not delve into the subtleties which some mathematicians regard as the essence of calculus. For our purpose that does not matter. The main thing we require is the vocabulary of calculus. If a reader is aware that ds/dt has some connexion with velocity and dy/dx with slope, that integration has to do with areas, and that e^x is related to compound interest and the way a population grows, this ability to use calculus as a language should go a long way towards enabling him to follow the themes of this book. Certainly, nowhere is any 'tricky' work, either in calculus or in algebra, invoked.

Further, in Chapter Six will be found a section headed 'Ersatz Calculus'. This shows how an electronic computer manages to reduce problems in calculus to problems in arithmetic. I do not think I would advocate the contents of this section as a first approach to calculus, but the section does give an account of calculus, which may serve as a reminder to readers who met calculus some time ago and have not had occasion to work with it recently.

The aim of the book then has been neither to drag calculus in nor to shut it out, but whenever a new mathematical idea is being described, which was obviously suggested by calculus, or has

among its natural applications some question in calculus, that fact has been duly noted. This may reduce the number of potential readers of the book, but it seems a lesser evil than producing new concepts out of the blue and leaving the reader in a state of perplexity as to their origin and function.

THE NEED FOR EVIDENCE

My hope, then, is that by the end of this book readers will not merely have met some new ideas, but will have seen at any rate some of the uses to which these ideas can be put. They can then judge for themselves whether they regard these ideas as important or not. It seems essential that readers should be provided with this kind of evidence, for what is important for one purpose may well be irrelevant for another. Indeed, the utmost confusion in discussion has been caused by the term 'modern mathematics' being used with a whole variety of different, and sometimes contradictory, meanings. Among these, we can distinguish the following:

Meaning A, the mathematical discoveries made since 1900, together with some earlier work that prepared the way for these discoveries. This meaning, I believe, was the one intended by those who first coined the slogan 'modern mathematics'.

Meaning B, the mathematics needed for the science and technology of today and tomorrow.

Meaning C, the changes in arithmetic and other parts of mathematics that are called for by the increasing availability of electronic computers, desk calculators, and other means of automatic computation.

Meaning D, any method of teaching mathematics, recently invented or currently popular.

Meaning E, a label a publisher puts on a book to make it sell, and without other justification.

Now all of these meanings – except the last – correspond to considerations that should affect the planning of mathematical education. We do wish, in planning a syllabus, to take account of all the mathematics that is known; we want our pupils to be able

to cope with the mathematical aspects of a scientific and technological age; we do not want to waste their time and effort on work that could be more efficiently done by a machine; we want them to have the best teaching possible. Satisfactory mathematical education can only be achieved by a proper balance between these considerations, and this is by no means easy to achieve in a world that is rapidly changing and in which there is no one competent to speak on all the departments of knowledge involved. A mathematician has to work very hard to learn even five per cent of the mathematics in existence today; he can hardly be expected to be well informed on the various sciences, on industry, and on teaching in schools. Other specialists are in a like plight. Teachers are confronted with the difficult task of drawing on the specialized knowledge of a variety of experts, and of welding their divergent ideas into a coherent whole.

This task sounds, and indeed is, extremely complex. But great harm is done by any approach which ignores this complexity. In some countries, at an early stage of the educational debate, mathematicians have been asked what they thought important, and it seems to have been assumed that their answers would automatically provide material relevant to the problems of industry and attractive to teach to young children. But the evidence for this mystical harmony is hard to find. Indeed, there is considerable evidence in the opposite direction. For specialists differ not only in what they know; they differ in their philosophies of life and in what they regard as important. To ignore this is to run the kind of risk you would take if you bought a car on the advice of a friend, and only afterwards discovered that, while you judged a car by the power of its engine and its mechanical performance, he judged it by its colour and artistic appearance.

Lest it be thought I exaggerate, I quote a recent article by Professor Dieudonné,* a leading mathematician and one who has contributed to the discussion on mathematical education. He complained that many people had a complete misconception of what he did. They thought of him as concerned with practical problems or using electronic computers. He did neither. He

* 'L'École française moderne des mathématiques', Jean Dieudonné, *Philosophia Mathematica*, vol. 1, no. 2 (1964).

went on to explain his view of mathematics today (the italics are Dieudonné's):

> The study of mathematical problems . . . leads us, little by little, to introduce . . . ideas much more abstract than those of number or shape . . . and which end up by having no longer *any interpretation in the world of the senses*. . . . These new notions pose in a natural way innumerable problems, to solve which we are led to introduce other concepts, even more abstract, in a swarm of an exuberant vitality, which, however, gets further and further from the origins of mathematics in Nature and so drives mathematicians more and more from the problems that physicists or engineers would put to them. . . . So one may say that in principle modern mathematics, for the most part, *does not have any utilitarian aim*, and that it constitutes an intellectual discipline, the 'utility' of which is *nil*. It can happen (as in the instances mentioned above) that abstract theories may one day find unsuspected 'applications'. All the same, it is *never* the idea of applications of this kind (which anyway are impossible to forecast) that guide the research mathematician, but rather the desire to advance the understanding of mathematical phenomena *as an end in itself*.
>
> No doubt, because of the historical origins of mathematics, many people find this viewpoint hard to accept, they always want mathematics to 'serve' something, and it seems shocking that mathematics should be merely a 'luxury' of civilization . . . mathematicians simply want people to recognize that they have the same right to independence that is given, for example, to astrophysicists, to palaeontologists, or to poets.

Dieudonné goes on to estimate that about eighty per cent of mathematicians today are completely uninterested in applications of mathematics.

Now I do not wish to deprive Professor Dieudonné of his independence and his freedom to continue doing the kind of mathematics he describes. It does seem legitimate, however, to point out that a technical college or a university with a technical bias should exercise some caution before accepting any advice that Dieudonné has given, or may in the future give, on the conduct of mathematical education, for it is clear that his values and purposes are somewhat different from theirs. Indeed, Dieudonné's testimony seems to indicate that a technical institution, concerned as it must be with utility, should turn its back on twentieth-century mathe-

matics and try instead to give a thorough knowledge of the mathematics developed in earlier ages, before mathematicians began to be driven 'more and more from the problems that physicists or engineers would put to them'.

Now there is undoubtedly much truth in Dieudonné's description of the present phase of mathematical research. Much of it is – so far as one can see – completely irrelevant to the needs of scientists, engineers, economists, managers, sociologists, and other users of mathematics. Yet it does seem that Dieudonné's picture is a little too sweeping. There do seem to be some strands in modern mathematics that are of practical as well as poetic interest. It does seem reasonable to suppose that some parts of twentieth-century mathematics will become as essential and as commonplace for the engineer of the future as seventeenth-century calculus has become for the engineer of today. The object of this book has been to identify and to present some of those topics in recent mathematics that are likely to be of value to people who are not professional mathematicians. Most of the topics chosen have also an intrinsic mathematical interest. Indeed, they illustrate one of the recurring themes of recent mathematics – that algebra, geometry, and calculus have a much wider scope than had formerly been imagined. In the past algebra, for example, was thought of as dealing with the properties of numbers. It is now recognized that all kinds of objects – operations, movements, etc. – have algebraic aspects. In the same way geometrical thinking and the processes of calculus can be applied much more widely than was ever imagined in past centuries. Examples of this will be found throughout the book.

These examples may help to throw light on one question which is controversial at the present time. There has been a reaction against 'routine manipulation' in algebra. Teachers do not like the idea of children slaving away at fifty exercises in order to produce mechanical slickness with the operations of algebra. In some places this reaction has gone so far that children do not know any of the traditional standard results in algebra. Such children will not be able to appreciate the situations in higher algebra that display an analogy with elementary algebra. It struck me, after several chapters of this book had been written, that these

chapters provided some evidence relative to this question. By examining them, one could see what background in elementary algebra was helpful as a foundation for work in modern algebra.

Going on to modern algebra is of course not the only reason for learning elementary algebra. Much of science still depends on the ability to use simple algebra as a language, intelligently and with understanding. This need is to be met, not by new mathematics, but by old mathematics, extremely well taught. This seems an appropriate place to mention one or two questions which lie outside the scope of this book, but which should be investigated and reported on if we are to have an adequate philosophy of mathematical education. This book is in the main concerned with twentieth-century mathematics that has grown out of earlier mathematics: it deals with branches that grow a certain way up the trunk. But there are also new shoots that have recently appeared from the ground near the foot of the tree. These are branches of mathematics that depend very little on earlier developments. Symbolic logic would perhaps be an example of this. It would be very useful to have a survey of such topics, covering not merely their mathematical content but estimating – so far as one can – their probable future impact on the life and work of mankind.

It would also be extremely useful to have some scientific forecast of the changes in education required by the growth of automation. Automation enables a machine to replace any human activity, physical or mental, that is capable of being reduced to a routine. Most present human activities are capable of such reduction, and much education is concerned with imparting routines; such education is clearly becoming obsolete. Automation will tend to concentrate human employment into occupations that call for specifically human attributes – originality, insight, judgement, initiative, understanding. Clearly the automated society will make great demands on the brilliant and the highly creative. A disturbing question is what such a society will need that can be supplied by a man or woman of average talents. It is not merely a matter of seeing that the material needs of all citizens are met. There will certainly be social disorders if a significant part of the population come to feel that they are only passengers and not contributing in any essential way to keeping things going. It

certainly seems probable that the minimum knowledge required for useful employment will steadily rise. This is already apparent in the United States, where there is unemployment among young people leaving secondary school with low educational qualifications and at the same time an unsatisfied demand for highly skilled technicians. Averting dislocation of this kind is a matter of human as well as industrial significance; it is something that most surely should be considered by teachers and others responsible for shaping education.

More knowledge is not only desirable because of questions of employment. Science affects all aspects of our life, from issues as large as those of the nuclear bomb to questions as intimate as the birth of thalidomide children. We have a greater power than ever before to interfere with the universe; it is desirable that this power should be widely understood, and used with knowledge and wisdom. The citizen of A.D. 2000 will certainly need much greater scientific background than the citizen of today; he may also need to know something of the techniques that operational research provides for arriving at rational decisions in a great variety of situations. In this general education, mathematics of some kind will certainly play a part.

It is to be hoped that studies will be made and published of all these issues.

ON COMPLEXITY AND DULLNESS

There is a stage in learning the piano where the only pieces you can play are the ones you do not consider worth hearing. The simple pieces strike you as boring; the interesting ones you find impossible. This difficulty occurs in many subjects. It was very noticeable in traditional algebra; one had to spend a long time on rather artificial and uninspiring questions before one's algebraic power was sufficient to solve any worth-while problem. It is of course the same too with modern algebra. I felt, in writing Chapters One and Two, that these were rather like the first act of a play. The characters have been introduced but they have not yet got tangled up in enough complications to be really exciting. Anyone who feels this should probably read these chapters rather

quickly. The uses of the results in them will increasingly appear as the book proceeds.

It might be worth while to mention the use made of footnotes. Some of them have obvious purposes – to supply the reference for a quotation, or some note on the historical origin of an idea. Some of them have a very definite purpose in relation to the question of difficulty. There is always a problem in writing a paperback as to how precise statements should be. The main aim of a paperback is to outline the leading ideas in a subject. So a writer starts off, and explains an idea in simple, rather general terms. Then he looks at what he has written, and decides it is not true, for in certain rather special circumstances, exactly the opposite would be the case. If he is not careful, he keeps on adding qualifying clauses, until the statement is about as readable as a legal document or an income-tax form, and the original purpose, a simple statement of a good general rule, has been completely lost. I have used footnotes as a way out of this. The simple statement appears in the text. The exceptional case, the objection that might trouble a particularly well-informed or critical reader, can be mentioned in a footnote. Some of the footnotes in Chapter Four have rather the function of an appendix; they outline calculations that are not essential for an understanding of the argument, and with which many readers will not wish to be bothered.

NOTATION

There are two notations for a function, an ancient and a modern one. If the ancient one is used, it upsets those who, with some effort and difficulty, have adjusted themselves to the modern one. If the modern one is used, it means that readers with a traditional background not merely have to learn new ideas but have to learn them, so to speak, in a foreign language. In my first draft of this book, I tried to find ways of wording statements about functions that would accord both with the old and the new usage. This led me into some rather tortuous sentences, and eventually I decided I had to come down on one side or the other. I chose the old. This was perhaps a hardy step, for feelings run high on this matter. Professor Hochschild, an eminent mathematician, reviewing a

book on Lie Algebras by Professor Jacobson, wrote sternly, 'On page 209 there is introduced a notational convention based on the barbarous principle of confusing a function with one of its values.' The 'barbarous principle' is the notation on which most of us were brought up. It is the notation used in Caunt's *Infinitesimal Calculus* and Lebesgue's *Lessons on Integration*, by which one speaks of 'the function x^2' or 'the function $f(x)$'. I would ask any readers who have been brought up on the more modern terminology to regard such phrases in this book as abbreviations for the slightly longer phrases which they would regard as correct.

My reasons for using the older terminology were as follows. First, I do not believe there is any real difference of thinking involved. If you ask a traditionalist to sketch the graph of x^2 and you make the same request, in slightly different terms, to a modernist, they both draw exactly the same parabola. The ideas are the same, the words are different.

Second, it seems to be the case that any inconveniences involved in the old notation appear only at a fairly advanced level. It is thus natural that a research worker like Hochschild should be strongly opposed to the old system, while a schoolboy learning to differentiate and integrate finds it perfectly acceptable. In this book, it is not until the last chapter, Chapter Ten, that these difficulties begin to be felt. Accordingly, it is in that chapter that I have discussed this question of notation. This discussion is in fact the last section of the book. Treating it at this stage allows us to see not only what the new notation is but also the reasons that led to its introduction.

A third reason is that anyone who wishes to read further in this kind of mathematics has to be bilingual anyway – I mean, they have to be able to cope both with the old and the new notations.

G. F. Simmons's exceptionally readable textbook *Introduction to Topology and Modern Analysis* (McGraw-Hill, 1963) uses the new notation, as also does the much harder book, Dieudonné, *Foundations of Modern Analysis*. The old notation is used in the advanced but beautifully written *Functional Analysis and Semi-Groups* by Einar Hille. It is also used in Smirnov's extensive work, particularly designed for workers in physical sciences *A Course in*

Higher Mathematics (English translation published by Pergamon).

If Lebesgue, Hille and Smirnov were able to think about modern mathematics with the help of the old notation, which is familiar to most readers, it seems to me that I am justified in using it in this book, which is no more than an introduction to the new ideas. But most certainly anyone who wishes to go further with the subject will need to master the new terminology and notation.

CHAPTER ONE

The Arithmetic of Space

ONE would not expect a young child to find any difficulty with either of the following questions – (1) What do you get if you add 3 cats and 1 dog to 1 cat and 2 dogs? (2) What is three times as much as 2 cats and 1 dog? These questions seem too simple to lead to any useful idea. Yet in fact they produce a simple but fruitful way of looking at geometry.

Let us label the collections involved in the first question.

$$\begin{array}{l} A = \text{3 cats and 1 dog} \\ B = \text{1 cat and 2 dogs} \\ \hline C = \text{4 cats and 3 dogs} \end{array}$$

The third line is found by adding the first two, so we write $C = A+B$. In Figure 1 this addition is illustrated on graph paper. The

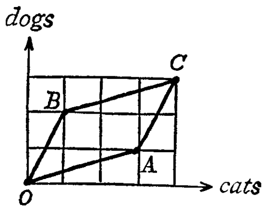


Figure 1

point A, with coordinates (3, 1), represents 3 cats and 1 dog; a similar explanation holds for B and C. It leaps to the eye that O, A, B, C are corners of a parallelogram. It can be tested by experiment that this result is not due to the particular numbers chosen. A cat-and-dog addition always corresponds to a parallelogram on the graph paper. (There are certain exceptional cases where the parallelogram is a 'thin' one. These arise when the points O, A, B are in line.)

The connexion between parallelograms and addition is of

course familiar to students of mechanics: parallelograms are used to add forces or velocities.

We now consider how our second question about multiplication looks on graph paper.

$$\begin{array}{rl} P = & \text{2 cats and 1 dog} \\ & \qquad\qquad\quad \times\, 3 \\ \hline R = & \text{6 cats and 3 dogs} \end{array}$$

This calculation shows $R = 3P$. The points O, P, and R are plotted in Figure 2. It will be seen that R lies on the line OP, but is three times as far from O as P.

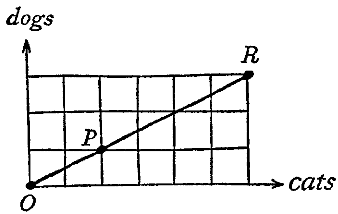


Figure 2

Multiplication is repeated addition, and it may help us to see the graphical significance of multiplication if we imagine ourselves starting with nothing and then adding 2 cats and 1 dog again and again. The calculation would go like this.

$$\begin{array}{rl}
\text{0 cat and 0 dog} & = O \\
+\text{2 cats and 1 dog} & \\ \hline
\text{2 cats and 1 dog} & = P \\
+\text{2 cats and 1 dog} & \\ \hline
\text{4 cats and 2 dogs} & = Q = 2P \\
+\text{2 cats and 1 dog} & \\ \hline
\text{6 cats and 3 dogs} & = R = 3P \\
+\text{2 cats and 1 dog} & \\ \hline
\text{8 cats and 4 dogs} & = S = 4P
\end{array}$$

Figure 3 shows the points O, P, Q, R, S. They are connected by something that looks like a staircase. In the arithmetic, at each stage we add the same thing, 2 cats and 1 dog. In the diagram, we go from each point to the next by taking the same step, 2 across and 1 up.

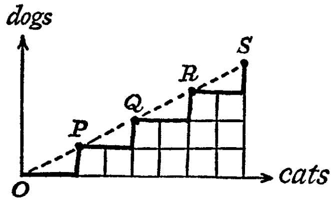


Figure 3

This idea of taking the same step is useful when we want to consider movements. In Figure 4 the heavy line $DEFG$ represents, say, a piece of wire lying on the paper. The point D^* is 2 across and 1 up from D; E^* is 2 across and 1 up from E; similarly, F^* from F and G^* from G. The heavy line $D^*E^*F^*G^*$ represents a new position the wire could take up. The arrows are meant to

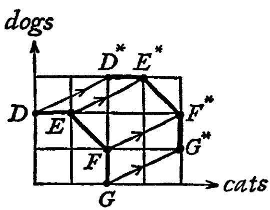


Figure 4

suggest this change, the wire moving from the old position $DEFG$ to the new position $D^*E^*F^*G^*$.

Such a change of position is called a *translation* (from the Latin *trans*, across, and *latus*, from *ferre*, to carry). In a translation every point is displaced the same distance in the same direction.

In our example, each arrow shows the effect of adding 2 cats and 1 dog, that is, the effect of adding P. We have $D^* = D+P$, $E^* = E+P$, and so on.

On page 20 we considered the effect of starting with nothing and then repeatedly adding P. Equally well, we could consider the effect of starting with any amount K of cats and dogs and repeatedly adding P. The effect would be as shown in Figure 5.

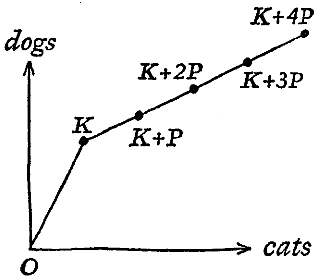


Figure 5

We now have a kind of arithmetic or algebra for describing positions in a plane, but does it really do us any good? To what sort of problem can it be applied? If we examine Figure 5 we may notice that the points labelled K, $K+P$, $K+2P$, and so on, are like the footprints of a man who walks steadily in a particular direction with even paces. For these points are in line and are evenly spaced. This suggests that our algebra may be particularly suitable for questions having to do with lines divided into equal parts.

So far the symbol P has stood for '2 cats and 1 dog'. It will be convenient at this stage to abandon that meaning and let P stand for any collection of cats and dogs that may be suitable for solving a problem. Our problems will be of the form: two points, K and L, are given; what pace P should we choose in order to get from K to L in a specified number of steps?

The simplest problem of this kind is: what is the formula for the mid-point M of KL? We suppose the points K and L are

specified in terms of so many cats and dogs, and that the mid-point M is to be found in the same form.

We shall land on the mid-point M if we walk from K to L in two paces. We hope to choose our pace P in such a way that K, $K+P$, $K+2P$ will coincide with K, M, and L. Where shall we find information to fix P? Not by looking at the first point (see Figure 6)

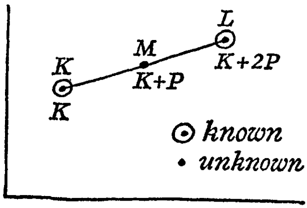


Figure 6

for it merely tells us $K = K$; not by looking at the second point, for M is still unknown and cannot help us to determine P (rather it is P that will lead us to M). However, the third point tells us $L = K+2P$, which is easily solved and gives $P = \frac{1}{2}L-\frac{1}{2}K$. Substituting, we have $M = K+P = K+(\frac{1}{2}L-\frac{1}{2}K) = \frac{1}{2}K+\frac{1}{2}L$. (In passing, it may be noted that this result shows M to be the average of K and L.)

For example, we might be asked to find the point midway between (2, 1) and (8, 3). We go over to animals: K = 2 cats and 1 dog; L = 8 cats and 3 dogs. So $\frac{1}{2}K$ = 1 cat and $\frac{1}{2}$ dog; $\frac{1}{2}L$ = 4 cats and $1\frac{1}{2}$ dogs. Hence $M = \frac{1}{2}K+\frac{1}{2}L$ = 5 cats and 2 dogs. We now return to our graph paper and announce (5, 2) as the mid-point M, which indeed it is easily seen to be.

We have here taken the liberty of talking about half dogs, and later we shall push poetic licence to the point of using such concepts as -3 cats. In fact, we are not taking this animal business too seriously. The reasons for using it are (1) to suggest that the ideas involved are simple, such as could be taught to young children, (2) to provide a situation in which addition and multiplication have natural meanings, (3) to provide convenient labels,

e.g. 'cat-and-dog addition', when later on we wish to refer briefly to the processes of this chapter.

Our formula for the mid-point puts us in a position to prove a well-known, but not very exciting, geometrical result; the diagonals of a parallelogram bisect each other. For simplicity, we suppose the parallelogram has one corner at the origin, O. If the other corners are A, B, C, as in Figure 7, we have $C = A+B$

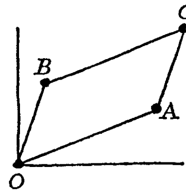


Figure 7

since, as was said on page 19, 'a cat-and-dog addition always corresponds to a parallelogram'. We want to show that the mid-point of OC coincides with the mid-point of AB. We prove this simply by calculating the positions of these mid-points, using the formula $M = \frac{1}{2}K+\frac{1}{2}L$ proved earlier.

The mid-point of AB is easy. It is $\frac{1}{2}A+\frac{1}{2}B$.

The mid-point of OC is $\frac{1}{2}O+\frac{1}{2}C$. Now O stands for zero cat and zero dog, that is, for nothing, so $\frac{1}{2}O$ also stands for nothing. So the mid-point of OC is simply $\frac{1}{2}C$. But we know $C = A+B$, so $\frac{1}{2}C = \frac{1}{2}A+\frac{1}{2}B$, and this, as we hoped, is the same as we had for the mid-point of AB. The theorem is proved.*

DIVIDING A LINE INTO ANY NUMBER OF PARTS

The argument we have used to find the mid-point is easily adapted to other, similar problems. Suppose, for example, we want a formula for the point S, three quarters of the way from K to L.

In this case, we want to reach L after four paces P from K. We want $L = K+4P$, and so we take $P = \frac{1}{4}(L-K)$. The values of

* The general result can be proved very similarly by considering the parallelogram with corners K, $K+A$, $K+B$, $K+A+B$.

Q, R, and S follow easily. In particular we find $S = K+3P = K+(\frac{3}{4}L-\frac{3}{4}K) = \frac{1}{4}K+\frac{3}{4}L$.

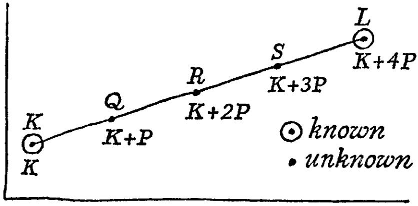


Figure 8

By examining this result, we can easily guess what the result would be if we used any other fraction instead of three quarters. We notice that three quarters appears as the coefficient of L. The coefficient of K is $\frac{1}{4}$, which is $1-\frac{3}{4}$.

Exercises

1. Guess formulas for the points one third of the way and two thirds of the way from K to L. Test your guesses by going through the full argument (forming and solving an equation for P).
2. Find the general formula for the point m/n of the way from K to L.

MEDIANS

Consider the following question. We are given the three points A, B, C (Figure 9); D is the mid-point of BC; find a formula for G, the point two thirds of the way from A to D.

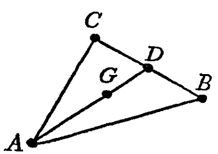


Figure 9

Given the results above (including the exercises), this is purely a routine calculation. The point G, being two thirds from A to D, must be $\frac{1}{3}A+\frac{2}{3}D$. Now D, being the mid-point of BC, must be $\frac{1}{2}B+\frac{1}{2}C$. Substituting and simplifying, we find $G = \frac{1}{3}A+\frac{2}{3}D = \frac{1}{3}A+\frac{2}{3}(\frac{1}{2}B+\frac{1}{2}C) = \frac{1}{3}A+\frac{1}{3}B+\frac{1}{3}C$.

This answer is symmetrical. It involves A, B, and C in exactly the same way, in spite of the fact that the question seemed to single A out for preferential treatment. So, if we had started at B and gone two thirds of the way towards E, the mid-point of AC, we should have landed on the same point G. A similar remark could be made, starting out from C. The point G in fact, as shown

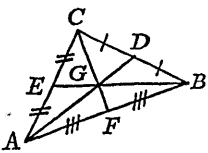


Figure 10

in Figure 10, lies on each of the medians AD, BE, and CF, and trisects each of them. G of course has some significance in mechanics as the centre of gravity of the triangle ABC, or the centre of gravity of equal masses placed at A, B, and C.

The demonstration just given that AD, BE, and CF have a common point is simpler than any proof available in Euclid's geometry, except perhaps by Ceva's Theorem. But how late Ceva's Theorem comes in Euclid!

We can see that any work based on our present methods is bound to be simple, for the only operations at our disposal are addition and multiplication by a number. However often we repeat these operations, we can never be led to really complicated algebraic expressions. We shall always be dealing with expressions of the first degree, such as occur in the exercises at the very beginning of an algebra book, 'add $2x+3y-z$ to $4x+5y+8z$', or 'multiply $5x+4y-3z$ by 7'.

This work with algebra we shall be able to interpret in geometrical terms. Our basic tool is the fact illustrated in Figure 5, that the points K, $K+P$, $K+2P$, $K+3P$ ... lie on a line and are evenly spaced. If three points U, V, W are specified (in cat-and-dog form) we can determine whether or not they lie in line. If they do, we can find the ratio of the distances UV and VW.

In *Prelude to Mathematics* the idea was used of discussing geometry with a disembodied spirit. The spirit was supposed to understand arithmetic. We will use the same device now. Suppose we are introducing some creature, with no geometrical experience, to geometry by the methods of this chapter. A point is defined as 'x cats and y dogs', or (x, y) for short. The results we have had in this chapter are used as definitions. The point midway between A and B is *defined* as $\frac{1}{2}A+\frac{1}{2}B$; the point $\frac{3}{4}$ of the way from A to B is defined as $\frac{1}{4}A+\frac{3}{4}B$; quite generally the point dividing AB in the ratio t to $1-t$ is defined as $(1-t)A+tB$, where t is supposed to lie between 0 and 1. (This definition is suggested by the result of the exercises on page 25.)

The spirit is now in a position to explore the plane by arithmetical methods. Suppose for example we ask it what it can discover about the points A, B, C, D, E, F, G, H, I where $A = (1, 1)$, $B = (2, 1)$, $C = (3, 1)$, $D = (1, 2)$, $E = (2, 2)$, $F = (3, 2)$, $G = (1, 3)$, $H = (2, 3)$, $I = (3, 3)$. It can report that B is the mid-point of AC, D of AG, H of GI, F of CI, while E is simultaneously the mid-point of AI, DF, GC, and HB. This we can see to be true by plotting the points on graph paper. To the spirit of course the statements are purely formal, arithmetical results; it has no graph paper and cannot imagine what graph paper is like. But the procedure we have given allows it to make calculations and produce statements in the language of geometry that will seem reasonable to us.

Does the procedure allow the spirit to develop the whole of Euclid's geometry? Anyone who has struggled with proofs in coordinate geometry will be convinced that the answer is, 'No.' The algebra of this chapter is much too simple for that. In fact the cat-and-dog procedure does not even mention several concepts that play a great part in Euclid, for instance the length of a line, or lines being perpendicular. In all our drawings so far the dog axis has been perpendicular to the cat axis, and the divisions along

the dog axis have been the same length as those along the cat axis. The reason for this was simple; that is the kind of graph paper most people are used to, and would have at hand if they wanted to experiment for themselves. But there is no justification in the nature of the mathematical topic for using this particular kind of graph paper. Why should dogs and cats be regarded as perpendicular, or as having the same length? In Figure 11 we see three different illustrations of the spirit's report on the points *ABCDE*

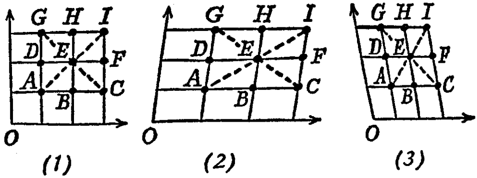


Figure 11

FGHI. In (1) the cat axis and the dog axis are perpendicular; in (2) and (3) they are not. In (1) the intervals on the cat axis are the same length as those on the dog axis; in (2) they are longer; in (3) they are shorter. Yet (1), (2), and (3) are equally good illustrations of the spirit's statements. In each of them B is midway from *A* to *C*, *E* is midway from *G* to *C*, and so forth.

ANGLES THAT HAVE NO SIZE

In certain mathematical theories we meet a difficulty. Two lines are mentioned; we innocently ask how big the angle between them is, and we are told that it has no size. Learners, not unnaturally, find this puzzling. There are two ways of dealing with this difficulty. I will call these the Axiomatic Viewpoint and the Erlanger Programme approach.

From the axiomatic viewpoint, it is supposed that we are given certain information and asked to work out all the consequences. The information given, together with its consequences, constitutes a mathematical subject. Our spirit has been given certain

information that enables it to make various geometrical statements. But none of these statements will enable the spirit to distinguish between situations (1), (2) and (3) in Figure 11. Exactly the same algebraic equations hold for these three diagrams. In each of them, for example, we have $A+B = F$, $E = 2A$, $B+D = I$. Any equation – of the kind we are considering – that holds for (1) will also hold for (2) and (3). Yet these figures differ in regard to angles and in the ratio of the lengths AB and AD. In a theory which is *solely concerned with the consequences the spirit could develop from his store of information*, it is as though these angles and ratios did not exist.

One point is worth mentioning before we pass to the other explanation of this difficulty. It *is* possible for the spirit to compare lengths when these lengths lie along parallel lines. In Figure 11 we have the equations $H = D+A$ and $I = A+2A$. This means, in our earlier phraseology, that you can get from D to H by taking one pace A, but you need two such paces to get from A to I. The spirit accordingly can recognize that the journey from A to I is twice as long as the journey from D to H. The spirit in fact has the material needed to prove a theorem of Euclid involved here; the line DH joins the mid-points of the sides GA and GI of the triangle AGI, so it must be parallel to and half as long as the base AI. The spirit, however, cannot attach any meaning to a comparison of lengths in different directions, for example, 'AD is half as long as AC'. In fact this last statement is true only for (1); it is false in (2) and (3).

THE ERLANGER PROGRAMME

The axiomatic viewpoint gives a perfectly clear definition of a mathematical subject, but a rather formal one. It supposes us confronted with a list of statements and invited to draw logical conclusions from them. But it is not clear how we should go about looking for such conclusions. The statements may completely fail to stimulate our imagination. We would like to have some way of seeing what we are doing. But here we run into a difficulty. If we make drawings, they may show too much; they may convey to us more than was in the original information.

The official name for the subject developed by our spirit is Affine Geometry.* In affine geometry, as we have seen, right angles do not exist at all and lengths can only be compared in special circumstances. But the world around us corresponds very closely to Euclid's geometry, and we see right angles and lengths everywhere. If we are to feel the meaning of affine geometry, we have to find some way of *destroying* part of the information given us by our senses.

We have already had a strong hint of how to do this in our discussion of Figure 11. Imagine our drawings made on the squared paper of diagram (1). However, suppose that, after we have made the drawings, someone is liable to come along and distort the paper in such a way that the squares become parallelograms as in (2) or (3). The only properties of our drawings that are relevant to affine geometry are those *that survive such distortion unchanged.* It can be proved that any property that does so survive can be recognized by our spirit and expressed in the algebra given to it. The distortions permitted consist of any combination of the following operations – (a) changing the scale of the 'cat' axis, (b) changing the scale of the 'dog' axis, (c) changing the directions of these axes. Another way of specifying the distortions is to say that a distortion is acceptable if points in a straight line are always sent to points in a straight line, and parallel lines are sent to parallel lines. In fact, affine geometry can be built up from the two concepts *straight line* and *parallel.*

From the axiomatic viewpoint, the store of information that leads to affine geometry is a part of the information that leads to Euclid's geometry. So every theorem that can be proved in affine geometry is necessarily a theorem in Euclidean geometry. But not all the theorems of Euclidean geometry are in affine geometry.

Affine geometry is much simpler than Euclid's geometry. Accordingly, if a theorem belongs to affine geometry, it pays to prove it by affine methods rather than the more cumbersome

* Euclidean geometry being the work of the great mathematician Euclid, one might suppose Affine geometry to be the creation of some mathematician called Aff. This is not the case. 'Affine' comes from the word 'affinity', to which Euler gave a special technical meaning in 1748 when he wrote on 'the similarity and affinity of curves'.

Euclidean theorems. We can easily test whether a theorem belongs to affine geometry or not by applying the distortions described above; if the theorem still remains true after the figure has been subjected to every acceptable distortion, then it belongs to affine geometry. All the theorems mentioned in this chapter, as being provable by cat-and-dog algebra, pass this test.

Already by the nineteenth century several geometries had been developed and recognized, for example Euclidean geometry, affine geometry, projective geometry, inversive geometry, and the non-euclidean geometries of Bólyai–Lobachevsky and Riemann. Considerable material was therefore available for a survey of possible geometries. A principle for the classification of geometries was enunciated by Felix Klein in the celebrated Erlanger Programme, the inaugural lecture Klein gave on becoming a professor at the University of Erlangen in 1872.

One of the main ideas in this lecture – though Klein did not put it this way! – was that you could tell which geometry you were in by seeing what you would object to people doing to your drawings. Geography could be regarded as the most restrictive geometry of all; you can alter neither distances nor directions without destroying the truth of your statements. In Euclid's geometry, you can be much more tolerant; we do not mind if a printer changes the scale of a drawing, or slides it across the page, or rotates it, or reverses it as in a mirror. The drawing will still illustrate the theorem just as well. In affine geometry all these things may still be done, and also the distortions we discussed earlier.* In projective geometry the freedom to distort is even greater; the figure may be replaced by a photograph of the figure taken from an oblique angle. Topology (or *analysis situs* as it was usually called in Klein's time) is the least restrictive of all. It allows the paper to be stretched or warped in any way you like, provided only that the paper is not torn. A property is a topological property

* In our cat-and-dog algebra, the point O stands for 'nothing', so it has a definite meaning, and we do not allow changes of origin. Strictly speaking, the algebra of this chapter corresponds to affine geometry, in which a particular point O has been singled out. It is therefore slightly more restrictive than affine geometry, since we cannot accept translations as allowable distortions.

only if it survives every such distortion unharmed. In topology, for example, we cannot distinguish between a triangle, a square, and a circle, for on a sufficiently elastic membrane any one of these may be deformed into any other.

There is much more to the Erlanger Programme than has been noted here. Our purpose has been simply to indicate the two ways in which we can think about a topic such as affine geometry – one way, logical, building up from the axioms, the other, pictorial, cutting down from Euclid's geometry by permitting distortions that will destroy irrelevant and unwanted properties of the picture. This latter device allows us to use pictures without being in danger of bootlegging into our thinking information not warranted by the axioms.

CHANGE OF AXES

We have seen that we are under no obligation to use perpendicular axes for our graph paper. Any two lines, provided they point in different directions, will do for axes. Two people, therefore, if asked to cover a plane with a network of parallelograms suitable for graph paper, might choose entirely different systems. How hard

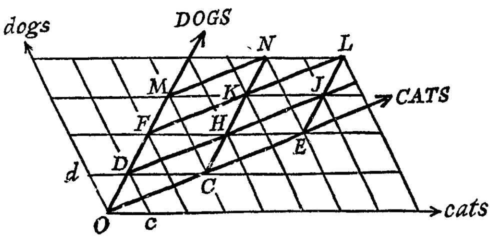


Figure 12

would it be to convert data, recorded in one system, into a form appropriate to the other? One might expect it to be very hard, but it is not. This problem also turns out to depend only on very elementary algebra.

In Figure 12 we see two systems of graph paper. One system has the axes marked *cats* and *dogs*. In this system, any point will be

specified as x *cats* and y *dogs*, or $xc+yd$ for short. The other system uses axes marked *CATS* and *DOGS*; any point will be specified as X *CATS* and Y *DOGS* or $XC+YD$ for short. Thus specifications in small letters refer to the first system, specifications in capitals to the second. We have perhaps been rather unfair to the first system in using capitals to label the points F, H, E, M, K, J, N, L for these points can be specified equally well in either system. On the other hand the points C and D are appropriately so marked, for they represent 1 *CAT* and 1 *DOG* respectively in the second system, while c and d represent 1 *cat* and 1 *dog* in the first system.

The table below shows a number of points with their specifications in the first and second system.

Point	In first system	In second system
C	$3c+d$	C
E	$6c+2d$	$2C$
D	$c+d$	D
F	$2c+2d$	$2D$
M	$3c+3d$	$3D$
H	$4c+2d$	$C+D$
J	$7c+3d$	$2C+D$

Comparing the specifications in this table, we notice certain things. D corresponds to $c+d$; $2D$ corresponds to $2c+2d$, just twice as much; $3D$ corresponds to $3c+3d$, three times as much. If this is not a chance coincidence, it means that we can work out the equivalents of $4D$, $5D$, $6D$, etc., without looking at the graph paper at all; for example, we expect $6D$ to be $6c+6d$. We notice the same effect with C and $2C$. In fact, the calculations seem to be simply those of the market place. If D exchanges for $c+d$, then $2D$ exchanges for $2c+2d$, $3D$ for $3c+3d$, and so on. How does this idea work when both C and D are involved? As C exchanges for $3c+d$ and D for $c+d$, we would expect $C+D$ to exchange for the sum $4c+2d$, and indeed it does. The entry for J also agrees with this method of calculation.

These results tie in with our geometrical picture of the meaning of addition and multiplication. Consider the point $4C+5D$. This is the point we should reach if we started at the origin, took four paces C and then five paces D. In the first system C is specified

by $3c+d$, so four paces C mean adding $12c+4d$; as D is specified by $c+d$, in the same way five paces D mean adding $5c+5d$. The final result is $17c+9d$.

The algebraic form of this calculation is simply a matter of substitution. $C = 3c+d$; $D = c+d$. Therefore $4C+5D = 4(3c+d)+5(c+d) = 17c+9d$.

What we have just done with $4C+5D$ could equally be done with any point $XC+YD$. We find $XC+YD = X(3c+d)+Y(c+d) = (3X+Y)c+(X+Y)d$. If this point is specified as $xc+yd$ in the first system, we must have the equations

$$x = 3X+Y \quad (1)$$
$$y = X+Y. \quad (2)$$

These equations tell us how to find (x, y), the coordinates of a point in the first system, when we know (X, Y), its coordinates in the second system.

Sometimes we may want to translate in the opposite direction. We may know (x, y) and want to find (X, Y). This is simply a matter of solving the simultaneous equations (1) and (2). We find

$$X = \tfrac{1}{2}x-\tfrac{1}{2}y \quad (3)$$
$$Y = -\tfrac{1}{2}x+1\tfrac{1}{2}y. \quad (4)$$

We are now in a position to translate statements from either system into the other. Such translation is very frequently needed. It often happens that we are forced to start a problem in one set of axes, and part way through we see that things would be very much simpler in some other system of axes, and so we change over. An example of this is the mechanics problem discussed at the beginning of Chapter Four.

Some of the older books on coordinate geometry give the impression that it is very difficult to work with oblique axes – that is to say, axes that are not perpendicular. They always start with perpendicular axes and trigonometry is involved whenever oblique axes come in. But both the systems in Figure 12 are drawn with oblique axes, and we have managed to translate from one to the other without even mentioning trigonometry. The work has involved nothing more advanced than linear expressions and, right at the end, solving a pair of simultaneous equations.

GENERALIZATION

This chapter began by considering collections of cats and dogs. The question naturally arises – why restrict ourselves to two kinds of animal? Why not consider calculations with three, four, five or indeed any number of animals?

Our procedure has been that of doing arithmetic or algebra, and then illustrating the operations pictorially.

It is fairly clear that the algebra will not be much different when we consider several kinds of animal. Adding 2 cats, 3 dogs, and 4 pigs to 5 cats, 6 dogs, and 7 pigs raises no new problem and presents no essentially new feature. It is quite different with the geometrical, pictorial aspect. When n, the number of animals, is three we can cope with the situation by going into three dimensions, the cat axis pointing (say) east, the dog axis north, and the pig axis up. But when n is four or more, our attempts at graphical illustration break down completely. The physical space in which we live has three dimensions and is completely unsuitable for illustrating additions involving more than three animals. By what, then, should our attitude to the cases where n is four or more be determined? – by the simplicity of the algebra, or by the absence of a physical model? Before we consider this question, let us look at the physical picture for $n = 3$.

THREE-D GRAPH PAPER

Most of us are capable of imagining solid objects only in a very vague manner. The coordinate geometry of three dimensions is regarded as a rather awe-inspiring subject, to be kept until late in the mathematical syllabus. This is a pity, for after all we do live in three dimensions; the great majority of the articles that we use or make occupy space in three dimensions, like an aeroplane or a motor-car, and are not spread out in a plane, like a carpet design or a printed circuit. Much three-dimensional coordinate work becomes simple, and can be studied by fairly young children, when it is done experimentally with the aid of an actual model. Such a model is easily made. By three-dimensional 'graph paper' we

understand a device that gives us a rapid means for measuring distances east, north, and up. Figure 13 illustrates a way of doing this. A piece of pegboard lies on the table. Upright pieces of dowelling can be stuck into the holes. In Figure 13, the point A is 3 inches east, 1 inch north, and 2 inches above the origin, O. The point A signifies 3 cats, 1 dog, and 2 pigs, or $3c+d+2p$ for short. In coordinate geometry it would be referred to as the point (3, 1, 2).

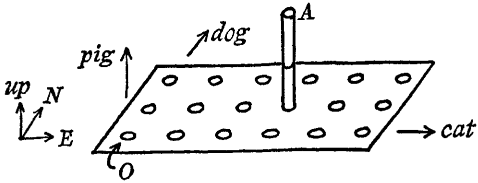


Figure 13

For practical use, certain modifications are needed. In order to make the dowelling stand firmly upright, it might be better to have a thick piece of wood, with holes bored in it, instead of pegboard for the base. Or alternatively one can bolt two pieces of pegboard together, with an air-space between, so that the dowelling passes through both boards and is securely held. Such details may be left to the maker of the model.

We now conduct an investigation along the lines of the argument for two dimensions. How, in the model, do we see the effect of adding? If $R = 3P$, how are P and R related? What is the effect of repeatedly taking the same pace P? Learners can answer these questions for themselves by experimenting with their pegboards, and may enjoy doing so. The answers will be found to show a great similarity to the results of our earlier work; these answers are given in the next few paragraphs.

Addition corresponds to parallelograms. Figure 14 shows the addition $C = A+B$ where:

$$\begin{array}{l} A = 4c+2d+p \\ B = c+3d+2p \\ \hline C = 5c+5d+3p. \end{array}$$

It is not easy to show it in the drawing, but the points O, A, B, C all lie in a plane and are the corners of a parallelogram. It would be

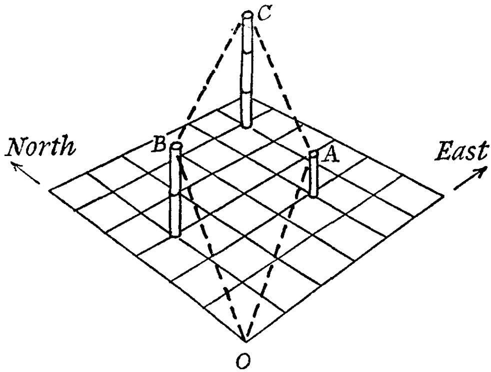


Figure 14

possible to cut a parallelogram from a flat piece of cardboard and place it in the position shown by dotted lines in Figure 14. It can be checked by further experiments that this result is not due to the choice of the particular numbers that occur in A and B.

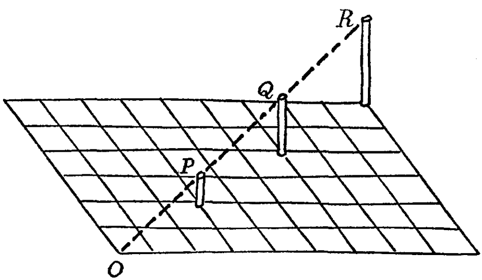


Figure 15

The effect of multiplication. Figure 15 illustrates the relation $R = 3P$, with $P = 3c+2d+p$ and $R = 9c+6d+3p$. Just as in two dimensions, we find O, P, R to be in line, with R three times as far from O as P.

Pacing. Figure 15 also shows $Q = 2P = 6c+4d+2p$, so that the points O, P, Q, R illustrate the effect of repeatedly adding P. These points lie at equal intervals on a line; our earlier image of pacing out the divisions of a line still works, but now we are pacing on a mountain side or sloping roof.

This result is important to us, for it was by means of pacing that we obtained our formulas for mid-point, and indeed for the division of a line in any ratio. *All these formulas hold, without any alteration whatever, in three dimensions.* For instance, if A, B, and C are any three points in space, it is still true that $G = \frac{1}{3}A + \frac{1}{3}B + \frac{1}{3}C$ represents the point where the medians of the triangle ABC meet.

As an example of this, Figure 16 shows a brick with one corner at the origin, O. D is the mid-point of BC and E of AC, so AD and BE are medians of the triangle ABC. They will meet at a point G somewhere inside the brick. By what has just been said, $G = \frac{1}{3}A + \frac{1}{3}B + \frac{1}{3}C$.

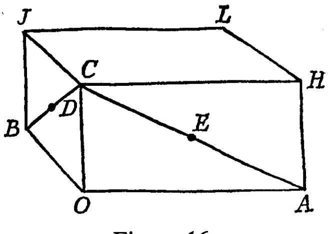


Figure 16

Now the end of the brick, $OBJC$, is a parallelogram (being in fact a rectangle), so $J = B+C$. If we take a pace equal to A from J we shall arrive at L, so $L = A+J = A+B+C$. Comparing the results for G and L, we see that L is exactly three times G; $L = 3G$. So G lies on the line OL and trisects it. Now G of course is in the plane of the triangle ABC, so it is the point where that plane meets the line OL. If we collect together the information we have found, we reach the following result; the plane ABC meets the line OL in a point G one third of the way from O to L, and this point G is

where the medians of the triangle ABC, meet. This is a formidable sounding result to have proved by simple arithmetic.

I have stated this result for 'a brick' because a brick is a familiar object, easily visualized and easily described; I do not have to use the awkward technical term 'rectangular parallelepiped'. But everything that was said earlier about our algebra and the distortions permitted in affine geometry remains in force. The argument nowhere depends on bricks having right angles at their corners. If the brick were squashed out of shape, but in such a way that all the faces remained parallelograms, the result would still hold true.

The same remark applies to our three-D graph paper. It is in no way essential that the axes point east, north, and up. The reasons for choosing perpendicular directions were essentially practical; it is not easy to obtain pegboard with holes arranged in parallelograms rather than squares, and with the holes bored in some oblique direction. Most people also find it easier to visualize three-dimensional figures in a rectangular framework; from the nursery we are brought up on rectangular bricks. But whatever the practical or psychological reasons for using perpendicular axes in illustrations, we should not lose sight of the fact that the mathematics of this chapter in no way requires the use, or even the existence, of right angles.

One general consideration emerges from the work we have just done, and it embodies one of the recurring themes of recent mathematics. In three dimensions we have obtained exactly the same formulas by exactly the same arguments as in two dimensions. This suggests the thought: could we not perhaps have found some way of developing these results without ever saying in how many dimensions we were? We would then have a theory holding for any number of dimensions. If we only recognize spaces of two or three dimensions, such a theory would save us repeating all our arguments twice, and thereby halve our work. But if, as we are going to consider in a moment, we find it possible to recognize the existence of spaces of four, five, six or n dimensions, the economy of thought is (literally) infinitely greater.

SPACE WITH FOUR OR MORE DIMENSIONS

Is it justifiable to speak of a space of n dimensions when n is bigger than three? As we saw earlier, our physical experiences have made us familiar with the geometry of the line (one dimension), the plane (two dimensions), and of space (three dimensions), but do not give us any direct way of visualizing four dimensions. What then is the status of the idea, space of four dimensions?

First of all, let us tear this question right away from all questions of physics. We are not concerned with the theory of relativity or whether time is in some sense actually the fourth dimension. We are not concerned with whether, in some other universe, or in some other part of this universe, there may exist creatures with the actual experience of living in six dimensions. We are concerned only with the mathematical soundness of the idea; we want to know whether correct thinking can be done using the idea of n dimensions – for in fact this idea is applied to some quite mundane, practical matters (in statistics for example) and we want to know whether we can rely on the conclusions reached with its help.

We can agree that the situation is perfectly clear on its algebraic side. If someone likes to consider collections of animals and carry out additions and multiplications, there is nothing whatever to prevent him from considering as many species of animals as he may wish.

Further, we have found a certain correspondence between operations with two or three animals and our physical experiences of flat and of solid objects. We can teach a disembodied spirit to translate an algebraic result, such as $A+B = C$, into a geometrical statement, $OACB$ is a parallelogram. What good does this do the spirit? None at all! A spirit has no experience of shapes; it finds the word 'parallelogram' meaningless and the whole exercise futile. It is we who benefit by the translation. We have spent many years in the physical universe; we are constantly observing the shapes and the movements of objects; our brains have built up an immense store of associations, which geometrical language evokes. For us the translation from algebra to geometry yields two benefits. On the one hand, it allows us to use the precise machinery

of algebra to calculate geometrical results that are not obvious to our visual imagination – for example, the property of the brick that we proved earlier. On the other hand, it allows us to picture formal results in the algebra. These pictures can give extra life to the equations of the algebra; by making the algebra more vivid, they help us to remember results, particularly those that correspond to rather obvious geometrical facts; these pictures also may help us to reason about the algebra, and to discover results that otherwise we would never have thought of.

The geometrical viewpoint is particularly helpful when changes of axes are involved. Suppose we have been working with one system of axes, and have established several results. Then it becomes desirable to change to another system. Can we still use these results or have we to test them all anew? It depends on the nature of the results. If they express what we might call *geometrical facts*, they will still hold in the new system; otherwise they may not. For example, if we prove that one point is midway between two others, we can be sure this result will not be affected by a change of axes. But when we have no geometrical interpretation of an equation, it is quite possible that that equation will no longer hold in new axes.

As an example of this, consider the points F, H, and E in Figure 12, and suppose we begin our work with the graph paper shown by thick lines. Then $F = 2D$, $H = C+D$, $E = 2C$, and we have $H = \frac{1}{2}E+\frac{1}{2}F$, so that H is the mid-point of EF. This geometrical fact must remain true if we decide to change to the graph paper with thin lines. Then we shall have $F = 2c+2a$, $H = 4c+2d$, $E = 6c+2d$, and in fact with these new symbols we do find, as we expected, that $H = \frac{1}{2}E+\frac{1}{2}F$. By contrast, consider that in the system with thick lines F has the coordinates (0, 2) while E has coordinates (2, 0). The coordinates of E are those of F in reverse order. However, when we go to the other system, F is specified as (2, 2) and E as (6, 2). The coordinates of E can no longer be obtained by reversing those of F. This property then was an accidental one, due to the particular axes used, and not valid after the change of axes.

There are some problems that drive us in the direction of studying spaces with four dimensions. In elementary algebra the graph of

$y = x^2$ helps us to understand the properties of that equation. The graph, of course, is in two dimensions, since two symbols, x and y, are involved. A little later, we meet complex numbers, involving $i = \sqrt{-1}$. If w and z are complex numbers, with $z = x+iy$ and $w = u+iv$, the equation $w = z^2$ now involves the four numbers, x, y, u, v; to graph it we would need space of four dimensions. The study of complex numbers would be very much easier if we lived in four dimensions and could actually draw such graphs. Not being able to see them, we still try to devise ways of thinking about them.

For these, and for many other reasons, we want to devise a geometrical way of talking about situations involving four or more numbers. So we treat ourselves in the way that earlier we treated the disembodied spirit; we provide certain rules for translating algebraic statements into geometrical ones. We have the advantage over the spirit that we do know what spaces of one, two, and three dimensions look like. The geometrical language therefore stimulates our imagination; it suggests analogies. Some of these analogies may be misleading; things may happen with four numbers that cannot happen with three. When we are in doubt, we go back to the algebra, to check on the correctness of our imagination. *So all questions of logic and proof are to be settled by algebra, or by arguing logically from geometrical statements which have themselves been proved by algebra.*

In a purely logical approach, then, such terms as 'parallel', 'in line', 'halfway between' would first appear as translations of situations in algebra. In the particular cases of one, two, and three dimensions, we would find that these geometrical statements agreed with their usual meanings, when we illustrated the algebra by means of graph paper. This would be an *experimental result.* For instance, right at the beginning of this chapter we saw that $C = A+B$ corresponded to the points O, A, B, C on the graph paper forming a parallelogram (in the everyday, physical sense of the words). I did not *prove* this result; I cannot prove it, and I do not need to. The correspondence between the algebra of cats and dogs and actual drawings on actual paper is a phenomenon of the real world; it can be established only by experiment, not by argument.

Of course you might say that Euclid's geometry gives a pretty accurate account of how real objects behave, and it is a (more or less) logical system. Could we not prove by Euclid's methods that $OACB$ must be a parallelogram? Yes, certainly we could. But Euclid's geometry is a vastly more complicated affair than the simple algebra of this chapter. Our aim is rather to show that, by giving our disembodied spirit the contents of this chapter, and a few extra instructions, he can prove all of Euclid's theorems. We hope to arrive at Euclid's geometry, rather than to set out from it.

Our sequence accordingly is as follows – begin with the algebra of cats and dogs; provide a dictionary for expressing its results in the language of affine geometry; verify by experiment that the theorems of affine geometry are useful for describing the real world; some time later, bring in some more assumptions (axioms) and with the help of these derive Euclid's geometry.

When, therefore, we speak of space of four or more dimensions we are not committing ourselves logically to any new belief. We are simply introducing picturesque language, which is found to be helpful in suggesting analogies between the algebra of n symbols and our everyday geometrical experience.

CHAPTER TWO

A Geometrical Dictionary

CHAPTER ONE concluded with the idea of a dictionary for translating from algebra to geometry. We will now give some details of such a dictionary. The procedure will be the same throughout. We will take correspondences between algebra and geometry that make sense in two and three dimensions, which we can imagine, and from these devise definitions to cover spaces of any number of dimensions, whether we can imagine these or not.

Straight line through the origin. We saw in Figure 3 that the points $O, P, 2P, 3P, 4P \ldots$ lay in line. If we consider other multiples of P, involving fractional, irrational, or negative numbers, it is not hard to discover that these also lie on the line; the negative multiples lie on the other side of the origin. Accordingly we define the line OP as consisting of all points of the form tP, where t is any real number whatever.

The letter t is chosen here because it suggests *time*. One can think of the line as being swept out by a moving point. At this moment ($t = 0$) it is at the origin; in three seconds time ($t = 3$) it will be at $3P$; five seconds ago ($t = -5$) it was at $-5P$. The line contains all the points where the moving point ever was or ever will be.

Imagine a pupil who starts badly but works hard, and whose marks improve in the extraordinarily regular way shown below.

	Arithmetic	English	Science	Geography	French
First week	0	0	0	0	0
Second week	10	3	7	9	2
Third week	20	6	14	18	4
Fourth week	30	9	21	27	6

The achievement of the pupil at any stage is shown by five marks, so we would need five dimensions to make a graph of his progress. If P represents his marks in the second week, $2P$ will represent his marks in the third week, for they are just twice as much, and $3P$

his marks in the fourth week. These points lie on a line through the origin in five dimensions.

Straight lines in general. A similar argument applied to the situation shown in Figure 5 leads us to define a straight line as consisting of all the points of the form $K+tP$, where K and P have fixed meanings, and t can be any real number whatever.

Segment of a line. A segment of a line means that part of the line that lies between two points A and B. We have already found a formula for the point lying between A and B, and dividing the line in the ratio t to $1-t$ (see page 27). This formula immediately gives us our definition; the segment AB consists of all points of the form $(1-t)A+tB$, where t varies from 0 to 1 only.

You may notice that this definition agrees with the previous definition of a straight line. For $(1-t)A+tB$ may be rewritten as $A+t(B-A)$. This is of the form $K+tP$. It corresponds to starting at A and taking the pace as $B-A$, which of course brings us to B after one pace. Thus we are at A at time $t = 0$ and at B at $t = 1$. At times between 0 and 1, naturally we are somewhere between A and B.

Exercises

We are in five dimensions. The letters c, d, e, f, g may be supposed to signify cats, dogs, elephants, frogs, and geese. Points are specified as follows:

$A = 5c+3d+e+f+6g$; $B = 3c+d+5e+f+4g$; $C = c+5d+3e+7f+2g$; $D = 2c+3d+4e+4f+3g$; $E = 3c+4d+2e+4f+4g$; $F = 4c+2d+3e+f+5g$; $G = 3c+3d+3e+3f+4g$.

1. Which point in the list above is the mid-point of AB?
2. If you started at D and kept taking paces P, where $P = c-e-f+g$, which points in the above list would you reach, and after what number of paces?
3. If you started at C and kept taking paces Q, where $Q = c-2d+e-3f+g$, which points would you reach, and after how many paces?
4. Is the mid-point of CG in the list above?
5. Is the point one third of the way from F to C in the list?
6. What pace would be needed to take you from C to E?
7. If you began at C and kept taking the pace calculated in question 6, through what points in the list would you pass?
8. What diagram do the points A, B, C, D, E, F, G together form?

INSIDE A TRIANGLE

We found earlier that the medians of the triangle ABC meet at the point G, where $G = \frac{1}{3}A+\frac{1}{3}B+\frac{1}{3}C$. G is a point inside the triangle. What about other points inside the triangle? How will they appear? Let us examine an example. Suppose Q is two thirds of the way from B to C, and R is three quarters of the way from A to Q (see Figure 17). What is the specification of R? We translate

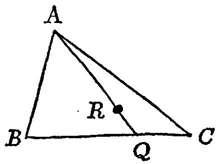


Figure 17

the geometrical statements above into algebra. R is three quarters of the way from A to Q means $R = \frac{1}{4}A+\frac{3}{4}Q$. Q is two thirds of the way from B to C means $Q = \frac{1}{3}B+\frac{2}{3}C$. Combining these results we have $R = \frac{1}{4}A+\frac{3}{4}(\frac{1}{3}B+\frac{2}{3}C) = \frac{1}{4}A+\frac{1}{4}B+\frac{1}{2}C$.

In the expressions for G and R certain fractions occur, and you may notice that in each case these fractions add up to 1. By experimenting with the methods just used, varying the fractions, one is led to believe that this always happens. It is not hard to show by algebra that it always does.

Accordingly we are led to the following definition of 'inside'; a point R is said to be inside the triangle ABC if $R = xA+yB+zC$, where x, y, z are positive numbers such that $x+y+z = 1$.

You may notice an analogy with our earlier definition of segment as consisting of all points of the form $(1-t)A+tB$. Here the fractions t and $1-t$ add up to 1. We could, if we liked, define the inside of the segment as consisting of all the points $xA+yB$, with x and y positive numbers such that $x+y = 1$.

INSIDE A TETRAHEDRON

Figure 18 shows a tetrahedron $ABCD$. G is the point where the medians of the triangle ABC meet, so $G = \frac{1}{3}A+\frac{1}{3}B+\frac{1}{3}C$. H is three quarters of the way from D to G. A brief calculation leads

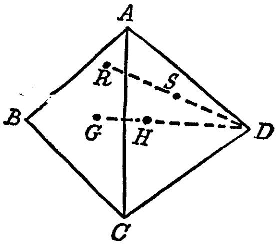


Figure 18

to the result that $H = \frac{1}{4}A+\frac{1}{4}B+\frac{1}{4}C+\frac{1}{4}D$. This point H must surely have geometrical properties in relation to the tetrahedron $ABCD$ similar to those G has in relation to the triangle ABC. For the moment, however, we are not concerned with these. We notice that H is inside the tetrahedron, and that the fractions in the formula for H add up to 1. Either by arithmetical experiments or by algebraic reasoning, we convince ourselves that if we take any point R in the face ABC and S anywhere inside the segment RD, we shall find S involving fractions whose sum is 1. Hence we define a point S as being inside the tetrahedron $ABCD$ if $S = xA+yB+zC+wD$ where x, y, z, w are positive numbers and $x+y+z+w = 1$.

For Figure 18 to make sense, the points A, B, C, D cannot lie in a plane. We naturally visualize this figure as lying in space of three dimensions. But the definition above, and our earlier definitions, still make sense if A, B, C, D lie in a space of, say, five dimensions. If, for example, $A = 5c+3d+e+f+6g$, $B = 3c+d+5e+f+4g$, $C = c+5d+3e+7f+2g$, $D = 7c+7d+7e+7f+4g$, we can still recognize $H = 4c+4d+4e+4f+4g$ as being $\frac{1}{4}A+\frac{1}{4}B+\frac{1}{4}C+\frac{1}{4}D$ and therefore a point inside the tetrahedron. We can still identify

$G = 3c+3d+3e+3f+4g$ as $\frac{1}{3}A+\frac{1}{3}B+\frac{1}{3}C$, and hence lying inside the face ABC. We can identify the mid-point of BC as $2c+3d+4e+4f+3g$ and check that G does lie inside the segment joining this point to A.

POINTS INSIDE A SIMPLEX

The figure for four points, A, B, C, D, has to be in three dimensions at least. For the next part of our argument, dealing with five points, A, B, C, D, E, we shall have to be in four dimensions at least; we have reached the stage where direct physical representation fails.

The analogy with our earlier work makes it pretty clear that we are going to define a point S as being inside $ABCDE$ if $S = xA+yB+zC+wD+vE$ where x, y, z, w, v are positive numbers and $x+y+z+w+v = 1$.

All the points that satisfy this condition fill a region of some kind. We need a name for this region, and shall call it a *simplex in four dimensions*. We could, if we liked, refer to the interior of a tetrahedron, a triangle, or a segment as a simplex in three, two, or one dimensions respectively. It should be evident how we would define a simplex in five, six, seven or n dimensions.

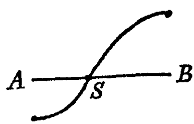


Figure 19

There are many simple arguments that we use in familiar geometrical situations. In Figure 19, we see a continuous graph, which is below sea level at A and above sea level at B. We can deduce that there must be a point S, somewhere between A and B, at which the graph is actually at sea level. This argument is frequently useful in solving equations numerically.

Again in Figure 20, if we are told that the point P is inside the triangle ABC while the point Q is outside, and that P and Q are

joined by a continuous curve, we can be sure that somewhere on this curve there is a point R lying on one of the sides of the triangle ABC.

It is sometimes not realized that part of recent mathematics is aimed, not at producing spectacular new results, but simply at adapting such familiar, simple, humdrum arguments to situations

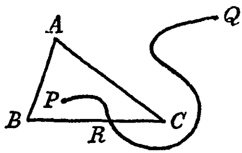


Figure 20

which may perhaps be dimly imagined, but where it is impossible to make the drawing that would show them to be obvious – as for example when the figure would require space with more than three dimensions.

A simplex can appear in quite natural and elementary problems. If we are mixing paint, we might use $\frac{1}{4}A+\frac{1}{4}B+\frac{1}{2}C$ to indicate that the mixture is $\frac{1}{4}$ red paint, $\frac{1}{4}$ blue paint, and $\frac{1}{2}$ yellow paint. The fractions automatically add up to 1. This simplex lies in two dimensions only, and can be shown as an actual triangle. But to represent the blending of five metals to form various alloys we should require a simplex in four dimensions. A dietetic study, showing the proportions in which seventeen articles of food were consumed, would require a simplex in sixteen dimensions.

Of course we cannot visualize a simplex in five or sixteen dimensions as completely as we can imagine a triangle. But even so, this geometrical description renders us part of the service that a graph does. It makes us aware that the regions in question are somewhat like a triangle, somewhat like a tetrahedron – only much more so! It tells us that they are limited in extent – they do not reach to infinity. They are all in one piece. They have pointed corners. This information is far from complete, but it can be suggestive.

If you had a large number of triangles, you could glue these together to make a great variety of shapes. You might make

diagrams in a plane, as in Figure 21 where a shape with a hole in the middle is made by joining together six triangles. You could also come out of the plane and make shapes that approximated to the surface of a sphere or a curtain ring. Similarly we could glue tetrahedra together to make more complicated solids. This device

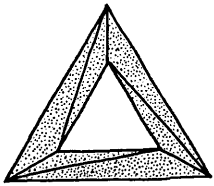


Figure 21

of joining simple objects together to make more complicated ones is one of the basic tools of combinatorial topology. The terminology corresponds to this practice; a *simplex* is the simple building block, the complicated object built from these is called a *complex*. Since a simplex gives a generalization of a triangle for any finite number of dimensions, mathematicians interested in combinatorial topology are not restricted to the spaces we can represent physically.

DROPPING RESTRICTIONS

We have seen that the interior of a triangle consists of all the points $xA+yB+zC$, where (1) x, y, z must be positive, (2) the sum $x+y+z = 1$. It is natural to wonder what difference it would make if we dropped either or both of these conditions.

One can explore this topic by taking simple cases, graphing them, and then seeking a logical explanation for what is observed. One would naturally start with the study of $xA+yB$. Here, to some extent, we already have the answer. Suppose first we keep the condition $x+y = 1$ but do not require x and y to be positive. We meet these conditions if we take any number whatever for x and set $y = 1-x$. So we have to consider all points of the form

$xA+(1-x)B$, which can also be written $B+x(A-B)$. This we have already met; it is the line we get by starting at B and taking all multiples (positive, negative, fractional) of $A-B$, the pace from B to A; it is the whole of the line AB.

Now suppose we drop both conditions and simply consider all points of the form $xA+yB$. This has an immediate graphical interpretation. The point xA lies on the line OA. The point yB lies on the line OB (see Figure 22). The sum $xA+yB$ is found by completing the parallelogram; it is represented by the point P. By suitable choice of x and y we can make P lie anywhere in the plane OAB.

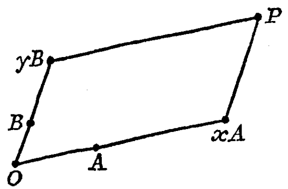


Figure 22

Definition of plane through origin. Accordingly, we define the plane OAB as consisting of all points of the form $xA+yB$. It is understood that O, A, and B are not in line.

If we were living in two dimensions only, the plane OAB would mean to us simply 'all the points there are'. But if we are working in three dimensions, this would no longer be so; we would be aware of plenty of points that were not in the plane OAB.

LINEAR SPACES

We now have two definitions – straight line OA, all points of the form xA; plane OAB, all points of the form $xA+yB$. In everyday life we stop here. The line has one dimension, the plane two, and we have no need to specify space of three dimensions. For us it is everything there is, the universe in which we move. But if we are going to imagine spaces of n dimensions we cannot stop in this

way. Just as a plane contains only some of the points in space of three dimensions, so a space of three dimensions can be made using only some of the points in four dimensions, and so on. There is an obvious way to continue the list above – linear space of three dimensions, $OABC$, all points of the form $xA+yB+zC$; linear space of four dimensions, $OABCD$, all points of the form $xA+yB+zC+wD$; etc., *ad lib.*

The definition of space of three dimensions just given agrees with our physical experience. If we take O, A, B, and C as they were in Figure 16 on page 38, OA points east, OB north and OC up. Any point in our universe can be reached by going a distance x to the east, y to the north, and z up.

Beyond this stage, physical experience ceases to help us. We cannot easily imagine our three-dimensional universe as merely a flat thing lying in space of four dimensions. (I ignore for the purposes of this illustration the theory that our space may be curved, and not a linear space at all.) For questions of this kind we have to rely on a mixture of algebra and reasoning by analogy.

A worked example. We are in space of four dimensions with basic symbols ('animals') c, d, e, f. Let $A = c-d$, $B = d-e$, $C = e-f$. Show that the point $c+d+e+f$ does not lie in the three-dimensional space $OABC$.

Solution. If it did, we would have $c+d+e+f = xA+yB+zC = x(c-d)+y(d-e)+z(e-f) = xc+(y-x)d+(z-y)e-zf$. Comparing amounts of c, d, e, and f, this means $x = 1$, $y-x = 1$, $z-y = 1$, $-z = 1$. The first three equations show $x = 1, y = 2$, $z = 3$ and this contradicts the fourth equation. So $c+d+e+f$ cannot be expressed in the form $xA+yB+zC$, that is, it does not lie in the space $OABC$. Q.E.D.

LINEAR DEPENDENCE

There is one point we still have to take care of. When we said that all the points $xA+yB$ formed a plane, we had to add the proviso that O, A, B were not in line. If OA points east and OB also points east, any combination $xA+yB$ will take you east. In our animal language, if A stands for 2 cats and B for 3 cats, then $xA+yB$ means $2x+3y$ cats. However we may choose x and y, we cannot

get off the cat axis. We are getting the points of a line, not those of a plane.

A similar effect can occur with spaces of any number of dimensions. Consider our familiar space of three dimensions, with c, d, and e representing points east, north, and up from the origin, O. Usually $xA+yB+zC$ would give us all the points of the space. But suppose $A = c$, $B = d$ and $C = c+d$. Adding A carries us east; adding B carries us north; adding C carries us north-east. All of these are horizontal journeys. However we may choose x, y, and z, we shall never get off the ground. The expression $xA+yB+zC$ conveys an illusory sense of freedom; it suggests that we have three ingredients A, B, and C to play about with. But A, B, and C are themselves mixtures of the two ingredients c and d.

Whenever we are disappointed in this way, we say that the ingredients are *linearly dependent*. Let us agree to say that the space consisting of all the points $xA+yB+zC$ is generated by A, B, and C, that the space consisting of all $xA+yB+zC+yD$ is generated by A, B, C, D, and so on. Then we can say that several points are linearly dependent when they generate a space of lower dimension than we would expect from their number, i.e. a space that could be generated by fewer points.

When a space is generated by several points, A, B, C . . ., we can imagine these points being given to us in turn. With A we expect to form the points xA of a line; when B arrives, we expect to extend this to the points $xA+yB$ of a plane; with the help of C we hope to construct a three-dimensional space, and so on. If each point. as it arrives, increases the dimension of the space by 1, we cannot fail to end up with a space of the expected number of dimensions.

If we get a space of lower dimension, it must mean that at least one point came in and failed to contribute anything new. This point must have itself been in the space already generated by its predecessors. We had an example of this earlier, with $A = c$, $B = d$, and $C = c+d$. The first two points, A and B, generated the plane of the ground. As C was also on the ground, it contributed nothing new. In the algebra, this is shown by the equation $C = A+B$, which shows that C is a combination of elements already present.

The breakdown can occur at any stage. We could have $A = 2c$, $B = 3c$, $C = d$. Here C does bring in something new, but B has already let the side down – and nothing can ever repair this! It is impossible for a single point C to bring in two new dimensions to make up for B's lapse. In the algebra, the failure of B to introduce any novelty is shown by the equation $B = (1\frac{1}{2})A$.

One might leap to the conclusion that the deficiency could not possibly be due to A, for A has no predecessors to plagiarize from. Yet failure is possible here. A may be O; it may represent no cat and no dog and no elephant; as an ingredient it may represent an empty packet. In the algebra this appears as the equation $A = O$.

The equations involved in our three examples were thus $A+B-C = O$, $3A-2B = O$ and $A = O$. In the middle equation, the coefficient of C is zero; in the last equation, the coefficients of both B and C are zero. But of course linear dependence would not be established by producing an equation in which *all* coefficients are zero, for such an equation does not say anything at all; any A, B, C whatever satisfy the equation $0A+0B+0C = O$.

In textbooks, this is often summed up in the formal definition of linear dependence; A, B, C are said to be linearly dependent if they satisfy an equation $pA+qB+rC = O$ where at least one of the numbers p, q, r is not zero. A similar definition of course applies to n points.

In the formal theory of linear dependence there are various logical points that have to be developed carefully. The theory is not particularly exciting since nothing unexpected happens. The careful reasoning simply confirms what you would guess to happen, by analogy with the familiar geometry of three dimensions. The concept of linear dependence however is an important one.

TESTING LINEAR DEPENDENCE

Suppose three points in space of five dimensions are specified as follows:

$$\begin{aligned} A &= c+2d+3e+4f+5g \\ B &= 6c+7d+8e+9f+10g \\ C &= 11c+12d+13e+14f+15g. \end{aligned}$$

Are these linearly dependent or not? Often it is necessary to answer a question of this type, and it can be done by the following method. We imagine the points produced in the order A, B, C. Obviously A is not O, so A does its job; it generates a line. What about B? B would fall down if it were merely a multiple of A. Since A starts off $c+\ldots$ and B starts off $6c+\ldots$, B could only be a multiple of A by being six times A. Then B would have to start with $6c+12d+\ldots$, but this does not agree with the $7d$ that actually occurs in B. So B is not a multiple of A; it has a different direction from A and the points $xA+yB$ fill a plane. Is C in this plane? We try to express C in the form $xA+yB$. *If* C is of this form, we can fix x and y by means of two equations, so we consider the first two terms only. As A begins $c+2d+\ldots$ and B begins $6c+7d+\ldots$, $xA+yB$ begins $(x+6y)c+(2x+7y)d+\ldots$. This will fit the beginning of C if $x+6y = 11$ and $2x+7y = 12$. These lead to $x = -1$ and $y = 2$. These are the only numbers that make C start correctly. If these do not give the whole of C correctly, then no combination of A and B will. As it turns out, it does work with these numbers; $-A+2B$ reproduces the whole of C perfectly. So the points A, B, C are linearly dependent. (As a matter of fact, B is the mid-point of AC.)

If we altered any one of the numbers occurring in C, this would make the points A, B, C linearly independent and they would generate a three-dimensional space.

A USEFUL RESULT

There is a certain rather obvious remark which turns out to be much more useful than it looks.

Suppose for example we are working in two dimensions, with basic symbols c and d, and that A, B, C are any three points in this space – perhaps $A = 3c+8d$, $B = 7c+5d$, $C = 27c+31d$. Now three points lying in a plane through O clearly cannot generate a space of three dimensions; being mixtures of only two ingredients, c and d, they can at most generate a plane. So they must be linearly dependent and so satisfy an equation $pA+qB+rC = O$, with p, q, r not all zero.

The particular points we have chosen in fact satisfy

$2A+3B-C = O$. But the argument works for *any* three points in two dimensions; they must be connected by an equation of this kind.

A similar argument holds if we are working in three dimensions. It is impossible for four points A, B, C, D in space of three dimensions to generate a space of four dimensions. So they must be linearly dependent, and hence connected by an equation $pA+qB+rC+sD = O$.

Quite generally, we can reach the conclusion that in a space of n dimensions, any $n+1$ points must be linearly dependent; they must be connected by an equation.

One could hardly call this a surprising conclusion. It is particularly natural when looked at from the geometrical point of view, and one might easily pass it over as of no special interest. Its value lies in allowing us, in varied circumstances, to show that something satisfies an equation.

Later in the book, we shall meet matrices, and see how this principle applies to them (see page 113). For the present, we consider an example involving only numbers. We want to prove that, if $t = 1+5\sqrt{2}+4\sqrt{3}+7\sqrt{6}$, then t satisfies an equation of the fourth degree – in which, of course, irrational numbers such as $\sqrt{2}$, $\sqrt{3}$, and $\sqrt{6}$ are not to appear. We write $c = 1$, $d = \sqrt{2}$, $e = \sqrt{3}$ and $f = \sqrt{6}$, so $t = c+5d+4e+7f$. Thus t is shown as a point in a certain space of four dimensions. Next we consider t^2. This is an unpleasant thing to work out by arithmetic, but fortunately we do not need to know what numbers appear in the answer. All we need to note is that t^2 also can be expressed in terms of 1, $\sqrt{2}$, $\sqrt{3}$, $\sqrt{6}$, and whole numbers; it also means a certain combination of c, d, e, and f. (In actual fact, $t^2 = 393+178\sqrt{2}+148\sqrt{3}+54\sqrt{6}$ so $t^2 = 393c+178d+148e+54f$.) The actual numbers do not matter in the least. All that matters is that t^2 also is a point in the same space of four dimensions. So are t^3 and t^4. We note also that 1 lies in this space, for $1 = c$. Thus we have found five points, $1, t, t^2, t^3, t^4$, in a space of only four dimensions. These points must be linearly dependent, i.e. there must be an equation $p.1+qt+rt^2+st^3+kt^4 = 0$ connecting them. It can be shown that the numbers p, q, r, s, k are rational; they do not involve $\sqrt{2}$, $\sqrt{3}$, $\sqrt{6}$, which indeed appear now only as d, e, f, in

the purely passive role played earlier by dogs, elephants, and frogs.

Our argument has thus shown that there must be *some* fourth-degree equation satisfied by t. This example illustrates both the power and the limitations of modern algebra. The argument above is simpler than any used in traditional algebra to prove this result. It avoids the lengthy computations of traditional algebra. It is excellent so long as we wish to show merely that there is *an* equation. If, however, we require to know *which particular equation*, we find ourselves forced to make detailed calculations. There are, though, some situations in which modern algebra suggests methods of actual computation that might not have occurred to us without its aid.

MODELS OF LINEAR SPACES

In this chapter and the previous one the mathematical theme has been rather monotonous; it has all been concerned with simple expressions such as $2c+5d+4e$. On the other hand the interpretations have been quite varied. Sometimes c is a cat, sometimes a point, sometimes one mark in arithmetic for a pupil, sometimes the number $\sqrt{2}$.

In each of these situations, we attached a meaning to the operation $+$. But these meanings differed considerably. When c stood for a cat or for $\sqrt{2}$, the meaning of $+$ was recognizable as *adding* either in the everyday or in the arithmetical sense of the word. But when A and B stood for points, $A+B$ signified the remaining corner of the parallelogram with three corners at O, A, and B. That $+$ should be associated with this operation is far from obvious.

Of course, the parallelogram definition of addition has long been used in mechanics and in those sciences, such as electricity and magnetism or aerodynamics, which are built on mechanics. In these subjects c and d are usually interpreted not as points A and B, but as arrows joining O to these points (Figure 23). So far as the calculations are concerned, it makes no difference whether we interpret c as the point A or the arrow OA. The arrow representation is often convenient. On page 21, Figure 4 showed the effect of adding the same thing, P, to D, E, F, and G. The shift

produced was shown by the arrows DD^*, EE^*, FF^*, and GG^*, all the same length and all pointing in the same direction. Using

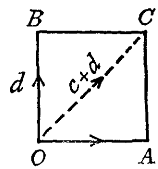


Figure 23

the arrow representation, we can add d to c by transferring the arrow d to the end of the arrow c (Figure 24). This saves us drawing the parallelogram; it simplifies the figure very much, particularly

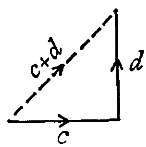


Figure 24

when, as in Figure 25, we have to add more than two things.

The idea that forces, velocities, accelerations, and various other physical entities could be added was already recognized in the

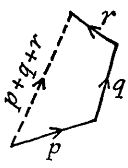


Figure 25

nineteenth century. The things represented by arrows were called *vectors* and the procedure for combining them was called *vector addition*. This idea then filtered across from mathematical physics

to pure mathematics, and a certain shift in emphasis and interpretation occurred. To the physicist and the applied mathematician, vectors were a recognizable class of actual objects. So to speak, you could tell whether a thing was a vector by looking at it. Velocity, for instance, was clearly suitable for drawing as a line with an arrow on it; mass was clearly unsuitable. So velocity was a vector, mass was not. But the pure mathematician was not interested in what physical objects, if any, his symbols might represent. He was interested in the mathematical pattern they made. So he came, in effect, to define vector space as *any collection of objects with which you can calculate by using the same formal rules as those that apply to vectors in physics*. In the language of Chapter One, vectors are things that can be combined *as if* they were collections of animals.

This brings in all sorts of things that look quite unlike forces and velocities. Compare the following calculations

$$\begin{array}{r} 2x^2+3x+4 \\ 5x^2+6x+7 \\ \hline 7x^2+9x+11 \end{array} \qquad\qquad \begin{array}{r} 2x^2+3x+4 \\ \times 5 \\ \hline 10x^2+15x+20 \end{array}$$

with these:

2 cats and 3 dogs and 4 pigs	2 cats and 3 dogs and 4 pigs
5 cats and 6 dogs and 7 pigs	×5
7 cats and 9 dogs and 11 pigs	10 cats and 15 dogs and 20 pigs

The same operations are being carried out in the two cases. Accordingly we may say that all possible quadratic expressions, ax^2+bx+c, form a vector space of three dimensions. Again compare these:

$$\begin{array}{l} 2\sin t+3\cos t \\ 4\sin t+5\cos t \\ \hline 6\sin t+8\cos t \end{array}$$

2 cats and 3 dogs	
4 cats and 5 dogs	
6 cats and 8 dogs.	

Once more, the calculations involved are identical. So we say

that all expressions of the form $a \sin t + b \cos t$ form a vector space of two dimensions.

Suppose we have a long wire tightly stretched as in Figure 26. Weights of P lb. and Q lb. are attached to the points B and C,

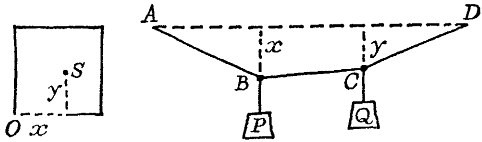


Figure 26

and cause these points to drop through distances x inches and y inches respectively. The position $ABCD$, to which the wire has been brought, can be specified by giving the two numbers x and y. The point S, with coordinates x, y, could therefore be used to indicate the position of the wire. This suggests that the possible

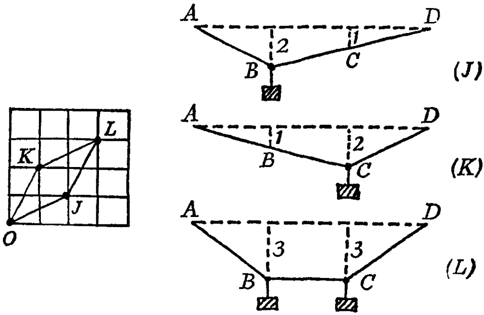


Figure 27

positions of the wire form a space of two dimensions. This would imply that we could talk about adding together two positions. Does this make sense? Figure 27 shows three points J, K, L, and

the situations to which they correspond. The information about these situations is collected in the following table.

Situation	P	Q	x	y
J	1	0	2	1
K	0	1	1	2
L	1	1	3	3

On the graph paper, O, J, K, and L form a parallelogram, so $L = J+K$. The table shows that this equation is physically meaningful. All the entries in row L can be got by adding together the corresponding numbers in rows J and K. In fact all possible situations can be deduced quite easily from this table. If we wanted to know what happened when 5 lb. was hung from B and 3 lb. from C, we would consider the situation $5J+3K$. We could extend the table like this:

Situation	P	Q	x	y
$5J$	5	0	10	5
$3K$	0	3	3	6
$5J+3K$	5	3	13	11

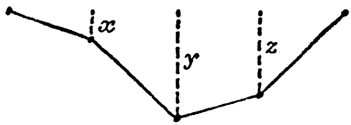


Figure 28

If we had three intermediate points on the string, as in Figure 28, we would need three numbers x, y, z to specify the position of the string. The state of the string could be shown by a point in three dimensions. From the pure mathematical viewpoint described above, we do not need to say that the position of the string can be *illustrated* in space of three dimensions. We can simply say that the possible positions of the string *constitute* a space of three dimensions.

If the string had seven points on it, at which displacements could occur, the positions would constitute a space of seven dimensions; with n points, a space of n dimensions.

In the figures and tables we had for $L = J+K$, x denoted the displacement of the point B. In situation J, x was 2; in situation K, x was 1; to find x in situation L, we simply add; $2+1 = 3$. Thus, to find the displacement of B in situation L, we need only

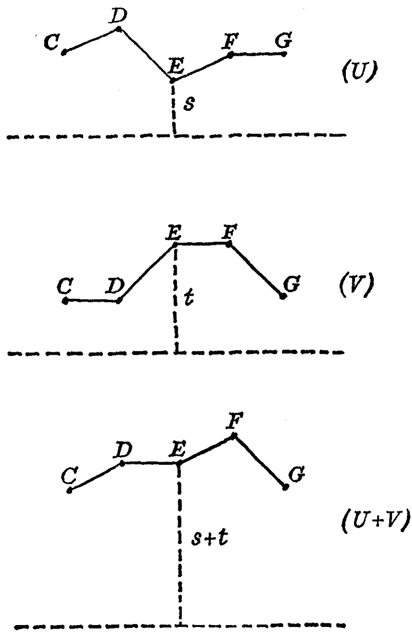


Figure 29

to know the displacements of B in situations J and K. We do not need to know anything about the displacements of any other point. This remains true however many points there may be on the string. In Figure 29 we see part of a string. In situation U, the point E has a displacement s; in situation V, a displacement t. So, in situation $U+V$, E must have the displacement $s+t$. By a similar

argument, we could plot the positions of the points C, D, F, G and any other points there might be on the graph for $U+V$.

In passing note that this is a very elementary and familiar procedure. If a firm sold boots and shoes, graph U might represent the sale of boots and graph V of shoes, at monthly intervals. The graph $U+V$ would then represent total sales. If s boots were sold in March and t shoes, the total sales in March would naturally be $s+t$ items of footwear. In a year, there would be twelve entries, so each graph would constitute a vector in twelve dimensions. Whether the manager would be any the better for knowing this is open to doubt.

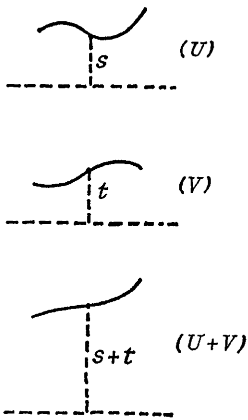


Figure 30

If we had a string with very many points on it indeed, our pictures would begin to look very much like smooth curves. This suggests that our procedure might be used with any graph whatever, not merely with graphs consisting of broken lines. And in fact there is no difficulty at all. Figure 30 is suggested by Figure 29, but the bends have been smoothed out. The construction is exactly the same as before. We read off s and t, the heights of E on graphs U and V; we then plot E on the graph $U+V$ at the height $s+t$.

This is the procedure we should have to follow if we were given two graphs, but not the equations or data used to draw them, and we wished to graph the sum.

If the graphs had simple formulas – for example, if U were $y = x^3$ and V were $y = x^2$ – then $U+V$ would be the graph we associate with the sum of the functions involved, namely, $y = x^3 + x^2$.

Quite generally, if U corresponded to $y = f(x)$ and V to $y = \phi(x)$, then $U+V$ would correspond to $y = f(x) + \phi(x)$.

We saw earlier that quadratic expressions such as ax^2+bx+c and trigonometric expressions such as $a \sin t + b \cos t$ could be added like cat-and-dog expressions and therefore qualified as vectors. We have now been led to a much wider conclusion; the same forms of calculation apply to any functions whatever, however wild and irregular. There need not be any simple formula; indeed, there need not be any formula at all. The functions could be specified by someone taking a pencil and making arbitrary free-hand graphs. These functions would still qualify as vectors. We could still do addition sums:

$$\begin{array}{r} 2f(x)+3\phi(x) \\ 4f(x)+5\phi(x) \\ \hline 6f(x)+8\phi(x) \end{array}$$

just as if $f(x)$ stood for cat and $\phi(x)$ for dog.

Functions then can be regarded as vectors. To most people this is surprising. We have kept functions and vectors in separate compartments in our heads. Functions we have associated with graphs such as $y = x^2$, vectors with arrows indicating wind speed 20 m.p.h. north-east. The considerations above show that the theory of vectors may enable us to gain information about functions.

Our first impression may be that this is a wonderful surprise. Our second impression may be one of disappointment. For this idea, however novel it may be, is not easy to cash in on. We would like to come down to brass tacks and show the value of the idea by producing definite results about functions and saying, 'See, I thought of these results because I knew functions could be re-

garded as vectors.' But at this stage our imagination does not get moving at all. For vectors are very simple things and there is not much to say about them – certainly nothing sensational. Our work with vectors has been based on the algebra of simple expressions such as $3c+4d$. That is not much mathematical machinery to have at our disposal. We shall have to bring in some more ideas before we get any very striking results.

This chapter concludes with a remark on terminology. *Vector space* and *linear space* are, for our purposes, synonyms. We can speak of x^2+2x+3 as being a particular vector in the vector space consisting of all quadratics, or we can say that it is a particular element (or point) in the linear space consisting of all quadratics. In principle, it would be possible to declare one of these terms obsolete. However, both terms are still current in mathematical literature and it is sometimes convenient to have them both available.

CHAPTER THREE

On Maps and Matrices

THERE are many situations in life in which one collection of objects will yield, produce, or be exchanged for some other collection. A ton of coal can be transformed in a gasworks into certain quantities of coke, ammonia, sulphur, benzene, and coal-tar. A coin put into an automatic machine will bring out a block of chocolate. Certain quantities of steel, rubber, copper, other materials, and human effort will produce a motor-car.

All of these processes are incapable of repetition. Having processed your coal in the gasworks, you cannot feed the coke and ammonia and so forth back into the gasworks and obtain some other range of products. The chocolate machine does not normally have any slot into which the chocolate can be put back and exchanged for some other commodity. A car factory does not take its new cars and put them through the assembly line again to produce some more complicated machine.

There are other processes capable of indefinite repetition. If you invest money, you can plough back your dividends into the same company and (with luck) get some more. If you breed animals, you may sell the offspring or you may keep them to breed from again. If a country has many factories, it can use them to make still more factories.

Both types of process – the once-for-all and the repeatable – have a certain mathematical pattern, which we will discuss in the simplified language of Chapter One, with animals instead of chemicals or car components. The applications of the resulting mathematical theory, incidentally, go far beyond the spheres of economics or production engineering.

ANIMAL BANKING SCHEMES

We imagine then a society in which all wealth is in the form of animals. If you put a cat into bank U, a year hence they will give

you a sheep. We write U; $c \to s$. This may be read 'in scheme U, a cat yields a sheep'. At bank V, if you take in a cat, they will give you a horse: V; $c \to h$. Bank W is more generous. For a cat they will give you a sheep and a horse: W ; $c \to s+h$. Evidently, scheme W is as good as schemes U and V put together. We indicate this by writing $W = U+V$.

We can thus add banking schemes. We can also multiply a banking scheme by a number. The scheme U is $c \to s$. By $3U$ we would understand a scheme three times as good as U, that is $c \to 3s$. In the same way, V being $c \to h$, the scheme $5V$ would denote $c \to 5h$. Finally, we might have a scheme as good as $3U$ and $5V$ together. We would write this as scheme $3U+5V$ and it would mean $c \to 3s+5h$.

If we suppose that banks accept cats only, and always pay out collections of sheep and horses, every banking scheme can be written in this way. If in some scheme a cat yields a sheep and b horses, the scheme is $aU+bV$. *The possible banking schemes form a linear space of two dimensions;* every possible scheme is of the form $aU+bV$.

The effect of a scheme can be shown geometrically. In Figure 31 we illustrate scheme W ; $c \to s+h$. The points O, L, M, N on

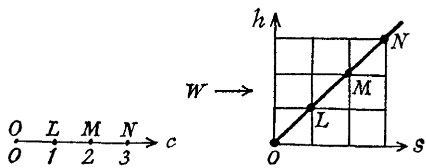


Figure 31

the line to the left represent possible investments – 0 cat, 1 cat, 2 cats, 3 cats. The corresponding points O, L, M, N at the right show the yields of these investments.

There is one thing this figure does not yet illustrate. We have just seen that the possible banking schemes form a space of two dimensions. The particular scheme W was arbitrarily chosen from this space. If we wanted to illustrate the situation completely, we

would need to envisage apparatus as shown in Figure 32. Here we have a panel in the middle. There are buttons on it corresponding to U, V, W and other banking schemes. The button W, drawn black, has been pressed, and this causes the points O, L, M, N to appear on the screen at the right. If another button, U for example, had been pressed, the yields corresponding to that button would have appeared.

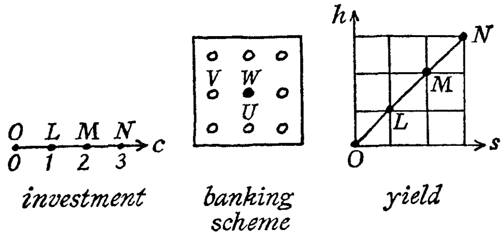


Figure 32

Instead of U ; $c \rightarrow s$ we may write $Uc = s$. 'The banking process U applied to c gives s.' In the same way V ; $c \rightarrow h$ may be written as $Vc = h$ and W ; $c \rightarrow s+h$ as $Wc = s+h$.

We have $W = U+V$, so we could also write $Wc = (U+V)c$. If for a moment we forgot what we were doing and relapsed into the habits of elementary algebra, we would be likely to continue the argument: $Wc = (U+V)c = Uc+Vc = s+h$. And this conclusion is in fact correct. The meanings of U, V, W, c, s, h are very different from those of elementary algebra, where symbols correspond to numbers. But the formal work above is exactly the same.

DIMENSIONS

In Figure 31 illustrating W ; $c \rightarrow s+h$, to the points O, L, M, N on the line at the left certain points O, L, M, N of the graph paper are made to correspond. We may say that the line is *mapped* into the graph paper. The operation W thus maps a space of one dimen-

sion (the cat line) into a space of two dimensions (the sheep and horse plane).

You might be tempted to think that there would be some restriction on what kind of space could be mapped into another. This is not so; given any two vector spaces, you can map one into the other by a 'banking scheme'.

If, for example, a banking scheme ran cat $\to$ sheep, dog $\to$ 2 sheep, elephant $\to$ 10 sheep, this would map a three-dimensional space into a one-dimensional space. For the investment requires three numbers to specify it, x cats, y dogs, z elephants; the yield requires only one number to specify it, n sheep. In fact, $n = x + 2y + 10z$. Of course, many points of the three-dimensional space would land on the same point of the one-dimensional space. A yield of 20 sheep might arise from an investment of 2 elephants, or of 10 dogs, or 20 cats, or from a combination such as 6 cats, 2 dogs, and 1 elephant. But there is no rule against this.

Every time your bill is made up in a shop, you illustrate a mapping from several dimensions to one dimension. The shop sells perhaps 100 different articles; you decide how many of each you want (choosing 'none' for most articles no doubt) so 100 numbers are needed to specify your purchase. Your selection identifies a point in space of 100 dimensions. But your bill is in terms of a single object, money; it can be measured as so many pence – one number only is involved.

If you were shortly due to travel by air, you might be concerned not only about the total cost of your purchase, but also about the total weight. Then, with each purchase specified by 100 numbers, you would associate two numbers, perhaps cost £12, weight 40 lb. You would be mapping from 100 to two dimensions.

The stretched string in Figure 26 can be regarded as giving a mapping from two dimensions to two dimensions. For we can choose the weights, P lb. and Q lb., that we hang at B and C. These will cause displacements, x inches and y inches at B and C; these displacements are in fact given by the equations:

$$x = 2P + Q$$
$$y = P + 2Q.$$

This mapping is shown in Figure 33.

Electrical illustrations of mappings are easily devised. In Figure 34a the voltage of the battery E may be chosen by us. It determines the sizes, x, y, of the two currents, so $E \to (x, y)$. This is a mapping from one dimension to two.

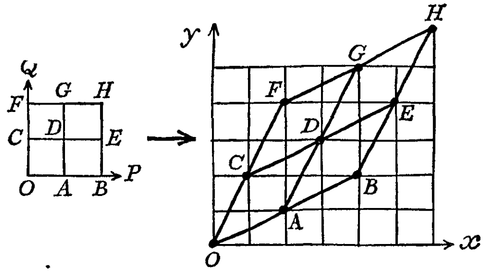


Figure 33

In Figure 34b also there are more wires than batteries. If we choose the two voltages E, F, these determine the four currents x, y, z, u. This is a mapping from two to four dimensions.

What particular mappings these will be depends on the resistances incorporated in the circuits. Here we have the same three

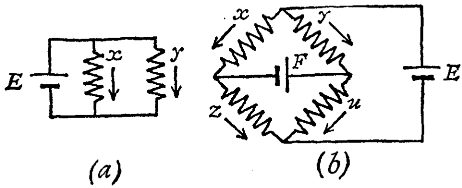


Figure 34

ingredients as those shown in Figure 32 on page 68. Voltages correspond to investment; choice of resistances corresponds to choice of banking scheme; the resulting currents correspond to yield.

COMBINING MAPPINGS

Imagine a situation such as the following. An oil refinery has two types of raw material available. A gallon of type (1) will yield a gallons of high-grade, b gallons of medium, and c gallons of low-grade petrol. A gallon of type (2) will yield d, e, and f gallons of high, medium, and low-grade petrol. Thus if x gallons of type (1) and y gallons of type (2) raw material are processed, there will result u, v, and w gallons of high, medium, and low-grade petrol, where

$$\left.\begin{aligned} u &= ax+dy \\ v &= bx+ey \\ w &= cx+fy \end{aligned}\right\} . \tag{1}$$

If the three grades of petrol sell at α, β, and γ pence a gallon respectively, the petrol produced will sell for t pence, where

$$t = \alpha u+\beta v+\gamma w. \tag{2}$$

We begin by choosing x and y; then x and y determine u, v, w; finally u, v, w determine t.

We have two mappings, P and S, corresponding to the operations of production and sale. P relates raw material used to petrol produced: P ; $(x, y) \rightarrow (u, v, w)$. S relates petrol produced to gross takings of money: S; $(u, v, w) \rightarrow t$.

We could write these as $t = S(u, v, w)$ and $(u, v, w) = P(x, y)$. It would then be natural to substitute and write $t = SP(x, y)$, so that SP would symbolize the overall process, raw material $\rightarrow$ money.

Note how our dimensions skip around here. We fix the raw material input by choosing two numbers, x, y. So the 'raw material' space is of two dimensions. The petrol produced is of three grades, so the output is specified by three numbers, u, v, w. The 'petrol produced' space has three dimensions. Finally, the money got is specified by a single number, t; the 'money' space has one dimension. These spaces, and the mappings connecting them, are shown in Figure 35.

Note also that the overall mapping, got by following the arrows

from raw material to money, is denoted by SP, not by PS. The money comes from the *sale* of the *product* of the raw material – not the product of the sale of the raw material. We also had a little algebraic argument earlier leading to the same notation. (Incidentally, this is a point on which there is the utmost confusion

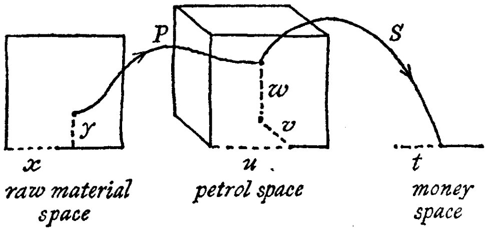


Figure 35

in mathematical literature. Whether to write SP or PS is a linguistic convention, and there is no general agreement which should be done.)

MULTIPLICATION OF MATRICES

The mappings P and S were specified by equations (1) and (2) above. By substituting from equations (1) into equation (2) we can get an equation linking t directly to x, y. This equation specifies the mapping $(x, y) \to t$. After multiplying out, we find it to be:

$$t = (\alpha a+\beta b+\gamma c)x+(\alpha d+\beta e+\gamma f)y. \tag{3}$$

A certain economy of writing is achieved by using what is known as matrix notation. This notation will prove in the future to have interesting consequences. For the present, however, it is simply a shorthand, obtained by omitting the letters u, v, w, x, y, t.

Thus we would specify P in matrix notation simply by writing:

$$P = \begin{pmatrix} a & d \\ b & e \\ c & f \end{pmatrix}. \tag{4}$$

This gives us the essence of equations (1). The six numbers, a,

b, c, d, e, f are entered in the positions they occupy in equations (1). If we wanted to, we could easily get back from the matrix statement (4) to the full, original form of equations (1).

In the same way, we extract the essence of equation (2) by writing:

$$S = (\alpha, \beta, \gamma). \tag{5}$$

From equation (3) we would have:

$$SP = (\alpha a+\beta b+\gamma c, \alpha d+\beta e+\gamma f). \tag{6}$$

Now we have specified SP in two different ways. We have specified it directly in equation (6). But we also know that SP is the combined effect of the mapping P specified by (4) and the mapping S specified by (5). We can make equation (6) above into a statement completely in matrix form, if we substitute for S from equation (5) and for P from equation (4). We shall then obtain:

$$(\alpha, \beta, \gamma)\begin{pmatrix} a & d \\ b & e \\ c & f \end{pmatrix} = (\alpha a+\beta b+\gamma c, \alpha d+\beta e+\gamma f). \tag{7}$$

From this equation we can obtain a mechanical rule for combining matrices. If you look at the first number on the right-hand side, $\alpha a+\beta b+\gamma c$, you will notice that the Latin letters a, b, c come from the first column of P, and they are multiplied by the corresponding Greek letters in the row for S. A similar observation can be made about the second number $\alpha d+\beta e+\gamma f$. It draws its Latin letters from the second column of P and multiplies them by the Greek letters in the row for S.

It is convenient to have such a mechanical rule. If we know two mappings in their matrix form, we can combine them by this rule without going to the trouble of writing out the equations in full and then substituting. But the mechanical rule is justified only by the fact that it produces the same answer that the full algebraic argument would do.

For instance, if someone were not clear from the above illustration by what rule to work out:

$$\begin{pmatrix} \alpha & \beta & \gamma \\ \delta & \epsilon & \zeta \end{pmatrix} \quad \begin{pmatrix} a & d \\ b & e \\ c & f \end{pmatrix}$$

he could find out this rule for himself. The first matrix expresses the essence of the two equations

$$\left.\begin{aligned} t &= \alpha u+\beta v+\gamma w \\ s &= \delta u+\epsilon v+\zeta w \end{aligned}\right\}. \tag{8}$$

The second matrix gives the essence of the equations

$$\left.\begin{aligned} u &= ax+dy \\ v &= bx+ey \\ w &= cx+fy \end{aligned}\right\}. \tag{9}$$

If we substitute for u, v, w from (9) into (8) we shall obtain equations giving t and s in terms of x and y. These equations in fact are

$$\begin{aligned} t &= (\alpha a+\beta b+\gamma c)x+(\alpha d+\beta e+\gamma f)y \\ s &= (\delta a+\epsilon b+\zeta c)x+(\delta d+\epsilon e+\zeta f)y. \end{aligned} \tag{10}$$

If we now erase x, y, t and s in equations (10) we shall be left with the four entries for the matrix we want. Incidentally this procedure also answers a question that often puzzles learners – what shape and size should the required matrix be? Equations (10) show that it should have two rows and two columns – a different shape from that of either of the matrices in the original question.

It is possible to write a question about combining matrices that makes complete nonsense, for example the following:

$$(\alpha, \beta, \gamma)\begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

The first matrix corresponds to the equation $t = \alpha u+\beta v+\gamma w$; the second to the equations:

$$\begin{aligned} u &= ax+cy \\ v &= bx+dy. \end{aligned}$$

Now we are in trouble when we try to substitute. We know what to put for u and v, but what about w? The problem is meaningless. Fortunately this also appears if we try to use our rule of running across the rows and down the columns. We start off all right, $\alpha a+\beta b+$, and then we are stuck. We have run out of Latin letters while there is still a Greek letter, γ, unused. We have nothing to go with γ. Our rule would be dangerous indeed if it led us to give a definite answer to a meaningless question.

One can also look at this matter geometrically. In Figure 35 we were able to form the mapping SP because P sent a point from the Raw Material space to the Petrol space, where S picked it up and sent it on to the Money space. Such a combination of mappings is only possible with a common space in the middle, and this imposes a restriction on the matrices representing the mappings. The output of the first operation must be suitable to form the input of the second. This was not so in the meaningless problem above, as may be seen from Figure 36. Operation

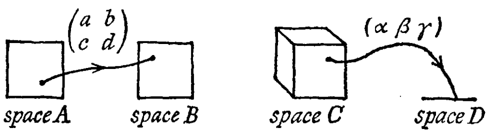


Figure 36

$\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ maps the two-dimensional space A into the two-dimensional space B. Operation $(\alpha,\ \beta,\ \gamma)$ maps the three-dimensional space C into the one-dimensional space D. There is no route from space A to space D.

LINEAR TRANSFORMATIONS OF A SPACE

The once-for-all operations that we have considered so far are all mappings from one space to another. The output is different in kind from the input. Cats go to sheep, coal to coke, pounds weight to inches, petrol to money. Such mappings have an importance, both in life and in mathematics, but as we have seen they are algebraically inconvenient. An operation can never be repeated, so we are never able to speak of PP or P^2. In some circumstances we can form combined operations, such as SP; in others, we cannot.

We have very much greater freedom when we are dealing with repeatable operations – those which map a space on to itself. It then becomes possible to combine any two operations and to form the powers of any operation, as we shall see below. Thus we can

use such expressions as SP, P^2, P^3, S^2P. We can still do everything that was done earlier in the chapter, so we are able to build expressions like $P^3+5P^2+7S^2P+2SP$. In fact we have a very large part of the machinery of elementary algebra at our disposal. In this chapter we shall see how this comes about. Later in the book we shall consider what use we can make of this machinery.

Earlier we considered schemes involving cat → sheep. Now we are concerned with schemes in which an investment of cats and dogs leads to a payment of cats and dogs which can be reinvested. Let us consider scheme K, in which $c \to c+d$ and $d \to d-c$. This may not make sense in banking terms, but it leads to an

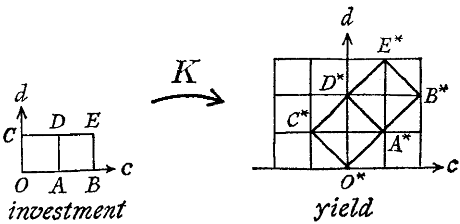


Figure 37

interesting geometrical picture. Figure 37 illustrates the effect of K as plotted on squared graph paper. The left-hand drawing shows the position of points $A, B, C, \ldots$ before the operation K is applied; the right-hand drawing shows the positions A^*, B^*, C^*, . . . to which these points go.

It will be noticed that the new points form a network very similar to that of the old points. This is no accident. Consider how we calculate the positions of A^*, B^*, C^*, D^*, and E^*. For A we simply look at scheme K, which says that for a cat (point A) you get a cat and a dog (point A^*). Then B corresponds to an investment of 2 cats, and this naturally yields twice as much. The yield B^* is thus twice the former yield, A^*; $B^* = 2A^*$. As we have seen, this means that B^* is in line with A^* from the origin O^*, but twice as far away. In the same way we can see that investments of 3 cats,

4 cats, etc. (points lying on the line OA) are bound to produce evenly spaced points on the line O^*A^*.

When we come to C^*, we simply read from scheme K that 1 dog (point C) goes to $d-c$ (point C^*). Next we come to D, representing an investment of a cat and a dog. Naturally the yield of this is got by adding together the yield of a cat (point A^*) and the yield of a dog (point C^*). Thus $D^* = A^*+C^*$. This means that O^*, A^*, C^*, and D^* form a parallelogram. Similarly, we see from the left-hand diagram that investment E is the sum of the investments B and C (for $OBEC$ is a parallelogram). The same must be true of the yields; we must have $E^* = B^*+C^*$ and this fixes E^* as the remaining corner of the parallelogram containing O^*, B^*, and C^*.

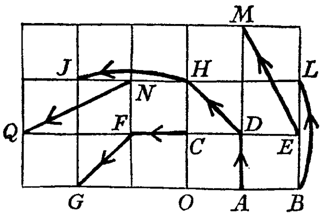


Figure 38

It will be seen that, once A^* and C^* have been plotted, all the remaining points could be got without further calculation, by drawing. This is the geometrical equivalent of the fact that when you know the yield of 1 cat and the yield of 1 dog, you know the yield of any number of cats and any number of dogs.

In Figure 37 we have drawn investment and yield on separate diagrams to aid clarity, but of course both diagrams represent the same space, the space of cats and dogs. The point A^*, for example, is the same point as D. Instead of two separate diagrams, we might use one piece of graph paper only, and draw an arrow from A to D, to show that operation K sends A to D. Since K sends every point somewhere, such an arrow would originate from every point of the plane. Some of these arrows are shown in Figure 38.

The arrows seem to organize themselves naturally into streams.

If we started at A and followed the arrows we should come in turn to D, H, J, and then to points not shown in this figure. Each step in this progression corresponds to the operation K. Investing A would yield D; reinvesting D would yield H; in turn, H would yield J, and so on. Thus we could write $D = KA$, $H = KD$, $J = KH$. From this we pass naturally to writing $H = KD = K(KA) = K^2A$; $J = KH = K^3A$. For J is what you get from A by applying the operation K thrice. Similarly, $G = K^2C$.

It is in fact quite easy to write down the banking scheme for K^2. As we have seen, K^2 sends A to H and C to G. As A represents c and C represents d, the operation K^2 makes $c \rightarrow 2d$, $d \rightarrow -2c$. This information is enough to tell us where every point goes.

With the graph paper we have used, K^2 has the geometrical meaning, 'turn through ninety degrees and double the scale'. However, as was discussed on page 28, there is no necessity to use squared paper to represent cat-and-dog space. The operations K, K^2, and K^3 could equally well be represented in the oblique graph papers shown in Figure 11. The geometrical interpretation of K^2 would then of course be different from that just given.

We should guard against a misconception that might arise from Figure 37. In this figure, $OADC$ and $ABED$ are squares, and the shapes $O^*\ A^*\ D^*\ C^*$ and $A^*\ B^*\ E^*\ D^*$ to which they go are also squares. This might suggest that banking schemes always send squares to squares, an idea that is false in two ways.

First, banking schemes are defined in terms of animal spaces, and in these spaces *we do not even know what squares are*. We have said nothing about cats being perpendicular to dogs, or having the same length as dogs. In short, in the language of Chapter One, we are still dealing with affine geometry, in which angles have no sizes, and lengths can only be compared when lines point in the same direction.

Second, if someone has decided to choose squared graph paper, from all the graph papers that are equally good for this work, it is still not true that every banking scheme maps squares to squares. In fact, to do so is exceptional. A simple example of a scheme that does not preserve squareness is $c \rightarrow c$, $d \rightarrow c+d$. Its effect is shown in Figure 39.

Later on in this book we shall consider what extra axioms have

to be brought in to stiffen up cat-and-dog space and make it into Euclidean space. We shall then be very much interested to pick out those banking schemes (called orthogonal transformations) which preserve the size and shape of squares. Until then, we shall be working within the confines of affine geometry and not con-

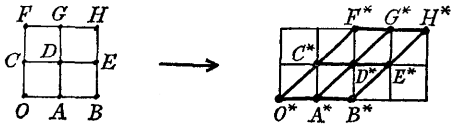


Figure 39

cerned with anything involving squares or right angles – at any rate so far as systematic exposition goes. From time to time it may be convenient incidentally to use some illustration with squared paper.

THE IDENTITY OPERATOR

The simplest operation of all is to do nothing, to leave things as they were. This operation is not admissible when we are mapping from one space to another, as with cat $\rightarrow$ sheep or petrol $\rightarrow$ money. No collection of sheep *is* a cat; no coin *is* a gallon of petrol. But when all our work is in a single space, there is no difficulty in leaving things as they are. This operation is known as the identity operation, and will be denoted by I. Its properties are reminiscent of multiplication by 1. If P represents any operation we have $IP = PI = P$. For IP means apply operation P and then leave things alone; with PI we leave things alone first and then apply P; either way the effect is P.

It may seem futile to introduce such a symbol. But consider the following question; if U is the operation $c \rightarrow d$, $d \rightarrow c$, what is U^2? Operation U changes each cat into a dog, and each dog into a cat. Clearly, if you apply this operation twice, you end up where you started; U^2 has the same final effect as inaction. So $U^2 = I$. Without the symbol I we would be unable to express this fact.

Exercise

What geometrical operation does U represent (on ordinary graph paper)?

MATRIX COMPUTATIONS

On page 72 we introduced matrix shorthand and showed how to calculate SP. Finding SP from S and P is usually called matrix multiplication. We have already explained what we mean by $U+V$ and $5U$ for linear transformations (banking schemes). If U and V were specified in matrix shorthand, it would be possible to work out, from these earlier explanations, the matrices for $U+V$ and $5U$. However, as such calculations are often needed, we will run through the argument here and give the results for reference, so that a routine procedure will be available.

Suppose U is a banking scheme in which x cats and y dogs $\to$ x^* cats and y^* dogs, where:

$$\begin{aligned} x^* &= ax+by \\ y^* &= cx+dy. \end{aligned}$$

Then U is represented by the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Similarly, suppose V to be the scheme with the matrix $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$.

Then $U+V$ is to mean the scheme that yields as much as schemes U and V together. Now x cats and y dogs with scheme U yield $ax+by$ cats and $cx+dy$ dogs; with scheme V they yield $\alpha x+\beta y$ cats and $\gamma x+\delta y$ dogs. So with scheme $U+V$, they must yield the sum of these, i.e., $(a+\alpha)x+(b+\beta)y$ cats and $(c+\gamma)x+(d+\delta)y$ dogs. Notice here how corresponding letters, a and α, b and β, and so on, join together as sums. From this specification of the scheme we can read off its matrix,

$$U+V = \begin{pmatrix} a+\alpha & b+\beta \\ c+\gamma & d+\delta \end{pmatrix}.$$

The rule for subtraction follows from that for addition, since $U-V$ means 'V and what make U?' We find:

$$U-V = \begin{pmatrix} a-\alpha & b-\beta \\ c-\gamma & d-\delta \end{pmatrix}.$$

The scheme $5U$ is that which yields five times as much as U. Thus $5U$ causes x cats and y dogs to yield $5ax+5by$ cats and $5cx+5dy$ dogs. By reading off the four numbers that appear here, we find the matrix result:

$$5U = \begin{pmatrix} 5a & 5b \\ 5c & 5d \end{pmatrix}.$$

A general formula could be found by replacing five by k in the argument and result just given.

The matrices U and V specify mappings of a two-dimensional space into itself. As we saw in Chapter One, linear spaces are not at all sensitive to the number of dimensions involved, and very similar results hold for the matrices that specify transformations in spaces of n dimensions. Such matrices have n rows and n columns. The rules for working with them can be readily guessed by analogy with those given above. Alternatively, you may like to take the arguments above, and rewrite them in the form they would have if three animals were involved.

We need to recognize the matrix form of the identity operator I. With this operator we associate the equations $x^* = x$, $y^* = y$. Comparing this with the general form $x^* = ax+by$, $y^* = cx+dy$, we see $a = 1$, $b = 0$, $c = 0$, $d = 1$. So we have

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

There is one other matrix we need to recognize. This is 0, the bankruptcy operator. Whatever you invest, you get nothing back; that is, $x^* = 0x+0y$, $y^* = 0x+0y$. Accordingly

$$0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

This operator, 0, has the properties we associate with zero. For any matrix U we find $U+0 = 0+U = U$ and $0U = U0 = 0$. These statements can be verified either by applying the rules of matrix computation or by thinking of the meaning of the banking schemes. For example, $U0$ means that you invest first in a firm that fails and gives you nothing; you then reinvest your return (which is nothing) in scheme U. Naturally you end with nothing.

In any number of dimensions the matrix for 0 contains noughts only.

In three dimensions the matrix $I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. We say it contains ones in the *main diagonal*, noughts elsewhere.

Having the symbol 0, we can write equations in the usual standard form of elementary algebra. For instance, the equation $U^2 = I$, which we met earlier, can be written $U^2 - I = 0$.

Exercises

1. If $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, $U = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, find $A+B$. Is it true or false that $A+B = 2I+U$?
2. Find A^2 and A^3. Guess A^n.
3. What is the difference between A^2+I and $2A$?
4. Simplify B^2-2B+I.
5. Let $W = I+U$. Find W.
6. Show $W^2 = 2W$.
7. If we substitute $I+U$ for W in the equation $W^2 = 2W$, and apply the rules of elementary algebra, we reach an equation containing U. What is this equation? Test by direct calculation whether the equation is true, i.e. whether U satisfies it or not.
8. Find AB and BA. Are they equal?
9. Find (1) $A^2+2AB+B^2$, (2) $A^2+AB+BA+B^2$. Are either of these equal to the square of $A+B$?
10. Find $I+2U+U^2$. Is it the same as the square of $I+U$?
11. Let $C = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Calculate $C^2-5C-2I$.
12. Let $D = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix}$. Calculate $D^2-3D-2I$.
13. Let $E = \begin{pmatrix} 1 & 5 \\ 2 & 7 \end{pmatrix}$. Show that for a certain number k, $E^2-8E = kI$. What is k?
14. Let $F = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}$. Does any simple equation connect F^2+7I and F?
15. Let $G = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$. Find G^2 and $2G+kI$. (The symbol k will of course appear in the second answer.) Is it possible to choose k so as to make G^2 and $2G+kI$ equal?

16. Let $H = \begin{pmatrix} 1 & 1 \\ 3 & 2 \end{pmatrix}$. Is it possible to find numbers q and k so that $H^2 = qH+kI$?
17. *Sustained investigation.* All the 2×2 matrices in the questions above have satisfied quadratic equations. Do all 2×2 matrices satisfy quadratic equations? Study this, either by experimenting with more particular examples, or by considering the general matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.
18. If $P = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $Q = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $R = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $S = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ find the matrix $P+2Q+3R+4S$.
19. Find $2P+Q+R+4S$.
20. Express the matrix G of question 15 by a combination of P, Q, R, S, similar to those used in questions 18 and 19. Do the same for I.
21. Two matrices occurring in question 1 are $A = P+Q+S$ and $B = P+R+S$. Adding these expressions would give $2P+Q+R+2S$. Does this agree with the value of $A+B$ found in question 1?
22. Can every 2×2 matrix be expressed in the form $aP+bQ+cR+dS$, where a, b, c, d are numbers?
23. Do 2×2 matrices constitute a linear space? If so, of how many dimensions?
24. What can be said about the space formed by all 3×3 matrices?
25. Multiply out $\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}$.

CHAPTER FOUR

On Hidden Simplicity

MATHEMATICS, according to Poincaré, is the art of giving the same name to different things. Mathematics thus leads to economy of thought, for we learn one mathematical pattern and then recognize it in many different situations.

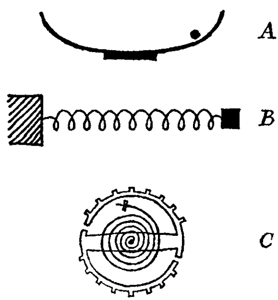


Figure 40

In Figure 40 we see three apparently different processes. In *A*, a mass oscillates near the bottom of a smooth curve: in *B*, a mass vibrates at the end of a spring: in *C*, we have the balance wheel of a watch. We know from experience that these three systems behave in very similar ways. Mathematically this appears in the formulas for kinetic energy, T, and potential energy, V. In each case $T = \frac{1}{2}mv^2$ and $V = \frac{1}{2}kx^2$: here v represents the rate of change of x. The form is the same, but the symbols have different meanings. In cases *A* and *B*, the symbol m represents the mass of the particle: in case *C*, it denotes moment of inertia. In cases *A* and *B*, the quantity x tells us (in somewhat different ways) how far the particle is from its position of equilibrium: in case *C*, it measures the angle through which the balance wheel has turned.

There is an interesting result in dynamics, according to which, provided a system is free from friction and satisfies certain very reasonable conditions, the behaviour of the system is completely determined by the formulas for kinetic and potential energy. If the three systems above have the same numbers for k and m, they will have identical formulas for T and V. This means that if they were started off suitably, they would remain in step with each other. By observing any one of them, we could tell what the others were doing. (It is assumed in each case that the oscillations are small.)

The first system, A, is particularly convenient to think about. The potential energy of a mass, acted on by gravity, is proportional to its height. So the curve in case A gives us a graph of the potential energy, $V = \frac{1}{2}kx^2$. When people go on the Big Dipper at a fair-ground, you can tell what their potential is. As they come slowly over the top of a hill, their potential energy is high. At the bottom of a dip, their potential energy is low. They shriek at the prospect of potential energy turning into kinetic.

The Big Dipper gives us a vivid picture of what potential and kinetic energy mean. We can often visualize a physical process by imagining an object careering about on an uneven landscape, which has been cunningly chosen so as to reproduce the potential energy of the system we are interested in.

We now apply this idea to the vibrating system shown in Figure

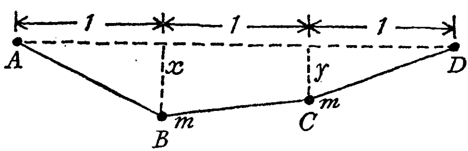


Figure 41

41. This system is chosen because it is simple to describe and to visualize, and its behaviour is typical of a very wide class of vibration problems of scientific interest or technical importance. The figure shows a string, tightly stretched between two pegs A and

D. Small masses m are attached at B and C. These masses vibrate up and down through small distances. The tension of the string is taken to be 1.

It can be shown* that the potential energy is given by $V = x^2 - xy + y^2$. So we proceed to construct a landscape that will give the same potential energy.

Imagine a sheet of graph paper lying on a flat table. At the point (x, y) we suppose a stick set up of height $x^2 - xy + y^2$. This is done at every point. The tops of these sticks form a surface. Imagine a copper bowl made so that it just rests on the top of these sticks. (In the best mathematical tradition, we ignore any inconvenient details. The thickness of the copper is assumed negligible.) A particle skates around in this bowl. When it is at position (x, y), its potential energy is $V = x^2 - xy + y^2$.

Thus we have reproduced the potential formula of the vibrating string system. Fortunately, it turns out that the formula for the kinetic energy is also reproduced.† Accordingly, if we observe how the particle slides in the bowl, this will tell us how the string vibrates.

Figure 42 is a contour map of the bowl. The contour lines are ellipses with equations $x^2 - xy + y^2 =$ constant. The bowl has a shape remotely reminiscent of a boat. It rises gently as you move north-east or south-west from the origin, much more sharply as you go north-west or south-east.

Wherever the particle is put on this bowl, it will feel itself pulled downhill. The downhill directions at various points are shown by the arrows in Figure 42. These arrows are perpendicular to the contour lines. A particle released from a point such as H would follow a complicated path, not easy to visualize.

However, two particularly simple motions can occur. If the particle were placed at G, it would slide straight towards O and

* The potential energy is due to the fact that the string is stretched in the position shown, and wants to get back to its original length. The lengths AB, BC, and CD can be found by Pythagoras' Theorem. We then use the fact that, for small k, $\sqrt{(1+k)}$ is approximately $1+\frac{1}{2}k$. This shows $AB = 1+\frac{1}{2}x^2$, $BC = 1+\frac{1}{2}(x-y)^2$, $CD = 1+\frac{1}{2}y^2$. So the string has been stretched by $\frac{1}{2}x^2+\frac{1}{2}(x-y)^2+\frac{1}{2}y^2$, which simplifies to x^2-xy+y^2.

† In fact, $T = \frac{1}{2}m(\dot{x}^2+\dot{y}^2)$, where $\dot{x} = dx/dt$ and $\dot{y} = dy/dt$.

come to rest at F; it would then return to G and repeat this oscillation indefinitely (no friction!).

In the same way, the particle could oscillate in the straight line

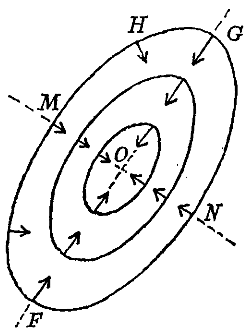


Figure 42

MON. This oscillation would be rather more rapid, since the bowl is more strongly curved in this direction than it is along GOF.

At this point a very welcome fact manifests itself. The motions along GOF and MON are not merely two isolated simple cases. By combining them, we can obtain *all the possible motions of the particle.*

Figure 43 shows how this could be done. Wire 1 is fastened rigidly to a metal piece which vibrates in the direction north-east to south-west. Wire 2 is similarly fastened to a piece that vibrates north-west and south-east. The first piece must be so arranged that its vibrations reproduce those of a particle oscillating on the path GOF. Similarly, the second piece must mimic the vibrations of a particle oscillating in the bowl along the path MON. Then the motion of the point R, where the two wires cross, will portray the motion of a particle sliding freely on the bowl.

It is interesting to see what the two specially simple motions – those along GOF and MON – mean in terms of the vibrating string. The line GOF has the equation $y = x$. A vibration along GOF is thus one in which x and y are always equal. Remembering

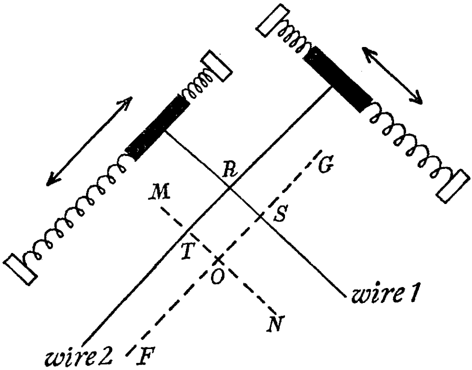


Figure 43

the meanings x and y have in Figure 41, we see that the string would pass through the stages shown in Figure 44.

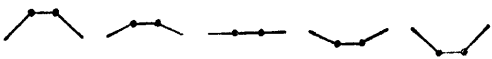


Figure 44

The line MON has the equation $y = -x$, and corresponds to a motion in which the displacements of B and C are always equal but opposite, as in Figure 45.

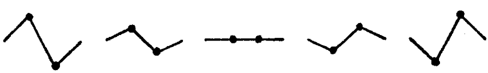


Figure 45

Anyone familiar with the theory of sound will see that these vibrations resemble the fundamental and first harmonic of a piano string.

If a string is set vibrating, the chances are that it will oscillate in such a way that a sensitive musical ear can hear *both* the funda-

mental and the first harmonic. Here again we have an example of *addition*; this complex sound could be produced by two tuning forks, vibrating simultaneously, one producing the fundamental note, the other the first harmonic. It would be reasonable to speak of this as the addition of the two sounds.

In fact, musicians played an essential part in developing the mathematical theory of vibration. The mathematicians fairly easily found the two simple solutions, corresponding to the fundamental and the first harmonic. It was only after the musicians had said they could hear both notes simultaneously that it occurred to the mathematicians to add their two simple solutions together and thus get the complicated solutions as well.*

The apparatus in Figure 43 is really a way of carrying out this addition. In that figure, $OSRT$ is always a parallelogram so – as in Chapter One – we may write $R = S+T$. As the wires move, S, the point where the first wire crosses GOF, dances backwards and forwards along GOF. The point T dances, at a higher frequency, backwards and forwards along MON. Since $R = S+T$ the motion of R can be regarded as a combination of the two dances. We are here finding an application of the process of vector addition developed in Chapter One.

Matrix operations also are involved here. When we displace the string, this naturally calls forces into operation that try to pull B and C back to their natural positions. There will be a force P tending to reduce x and a force Q tending to reduce y (see Figure 46a). Since the particle in the bowl gives a perfect dynamical picture of what the string does, it must be possible to discover corresponding forces P and Q in that system. They act in fact as shown in Figure 46b, P pulling the particle west and Q pulling it south. It is the combined effect of P and Q that is shown by the 'downhill' arrows in Figure 42.

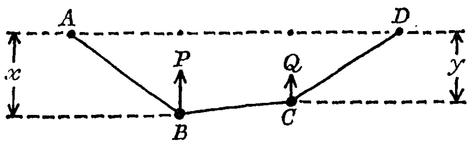
Actually the forces P and Q are given by the equations:

$$P = 2x - y$$
$$Q = -x + 2y.$$

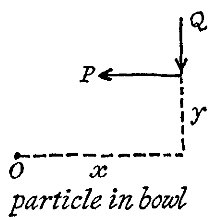
*See *Jahresbericht der Deutschen Mathematiker Vereinigung*, volume 10, II, 1 (1901–3), page 5. D. Bernoulli, around 1730–50, first saw the connexion. The other mathematicians refused to accept his views. He was right, they were all wrong.

These are linear equations, and so could be written in matrix shorthand.

A lesson that emerges clearly from Figure 43 is that our axes were not well chosen. In Figure 42, x is measured east and y north. But we never hear of east and north again. These directions have no significance. What we do hear about, again and again, are the lines *FOG* and *NOM*, pointing north-east and north-west.



(a)



(b)

Figure 46

Now of course we had to start with x and y, the displacements of B and C. These were the natural variables in which to express the forces P, Q, and the potential energy V. We could not avoid x and y so long as we were *stating* the problem. But as soon as we went over to *solving* the problem, it became desirable to bring in new axes, pointing north-east and north-west.

I do not wish to complete this problem in detail as it would involve more discussion of forces, accelerations, and differential

equations than is convenient at this point.* We will concentrate our attention on the central feature, the linear equations above that give P and Q in terms of x and y. We have the hint that these equations should become more tractable if we introduce new axes, FOG and NOM.

STUDY OF A TRANSFORMATION

At this stage it is convenient to drop the symbols P, Q and replace them by x^*, y^*. Accordingly, we shall study the transformation $(x, y) \to (x^*, y^*)$ where:

$$\left.\begin{aligned} x^* &= 2x-y \\ y^* &= -x+2y \end{aligned}\right\} . \qquad (1)$$

As these equations stand, they do not convey any idea to our minds; they do not enable us to see what the transformation does. Let us see if this situation improves when we bring in axes to the north-east and north-west. In Figure 47 the vectors C and D have these directions. The old axes correspond to the vectors c and d.

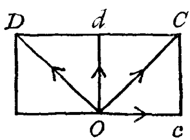


Figure 47

What does the transformation do to C and D? In our old axes, C is the point $x = 1, y = 1$. Putting these values in equations (1) we find $x^* = 1$, $y^* = 1$. The transformation leaves C exactly where it was. If you invest C, you simply get C back: $C \to C$. Now for D; D is the point $x = -1$, $y = 1$, and it leads to $x^* = -3, y^* = 3$. Thus $D = -c+d$, and the transformation sends it to $D^* = -3c+3d$. We see at once that D^* is the same as $3D$.

*Nothing formidable is involved. The equations of motion are $\ddot{x} = -2x+y$, $\ddot{y} = x-2y$. To bring in new axes we put $x = u+v$, $y = u-v$. This substitution leads to the equations $\ddot{u} = -u$, $\ddot{v} = -3v$, whence $u = A\cos t+B\sin t$, $v = C\cos(t\sqrt{3})+K\sin(t\sqrt{3})$.

The transformation can thus be specified as $C \to C$, $D \to 3D$. An investment of $XC + YD$ would thus lead to a return of $XC + 3YD$. Reading off the amounts of C and D in the return, we see:

$$\left.\begin{aligned} X^* &= X \\ Y^* &= \quad 3Y \end{aligned}\right\} \qquad (2)$$

These equations are simpler than those we started with, and they have a very evident geometrical meaning. $X^* = X$ means that we make no change in the first coordinate; $Y^* = 3Y$ means that we enlarge the second coordinate three times. This mapping is

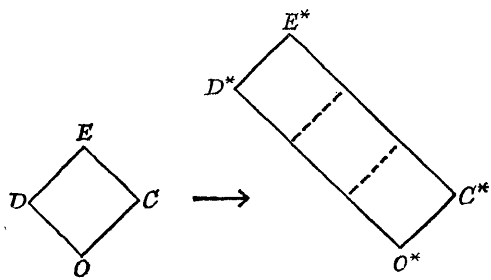


Figure 48

illustrated in Figure 48. It is a very simple, concertina-like operation. But how well it was concealed! This mapping was suggested by the equations $P = 2x - y$, $Q = -x + 2y$ linking displacements and forces for the stretched string. There was nothing in the problem of the vibrating string to suggest that it embodied so simple a transformation. The simplicity was entirely obscured by the unsuitable choice of axes.

DIAGONAL FORM

In the last section, equations (1) corresponded to the matrix $\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$. By a change of axes, we obtained equations (2), which correspond to the matrix $\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$. This last matrix has

noughts everywhere except in the main diagonal; it is said to be in *diagonal form.*

When a matrix is in diagonal form, its geometrical meaning is very evident. Consider, for example, the matrix $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$. This corresponds to the equations $x^* = 2x$, $y^* = 3y$. In Figure 49 this transformation sends P to P^*. The effect of the transformation is to double the scale on the x-axis and to treble it on the y-axis. One

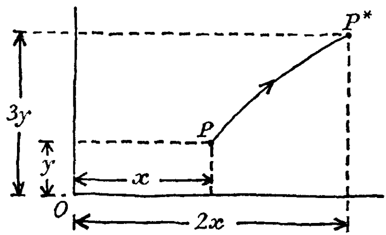


Figure 49

can imagine the effect as the stretching of pictures drawn on an elastic sheet.

Some exercises are given below, in which a change of axes leads to a matrix in diagonal form. All of these can be dealt with by the method of the previous section.

Exercises

1. In the section *Study of a transformation* we introduced new axes based on $C = c+d$, $D = -c+d$. Find how the following transformations appear in the new system:
   (a) $x^* = y$, $y^* = x$;
   (*b*) $x^* = 3x-y$, $y^* = -x+3y$;
   (c) $x^* = x+y$, $y^* = x+y$.
2. The transformation in 1(a) above corresponds to one of the matrices specified in the exercises at the end of Chapter Three; to which one? To which does 1(c) correspond?

EIGENVALUES AND EIGENVECTORS

We have just considered the transformation $x^* = 2x$, $y^* = 3y$ and

its geometrical meaning. Figure 50 shows various points A, B, C . . . and the positions A^*, B^*, C^* . . . to which this transformation

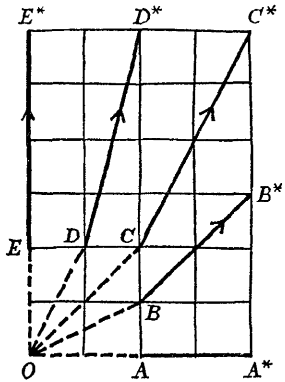


Figure 50

sends them. At each of the points B, C, D a bend is apparent. For example, the points O, B, B^* are not in line; the direction OB^* lies closer to the north than the direction OB. If we choose any vector v, and $v \rightarrow v^*$, *as a rule* we shall find that v and v^* have different directions. There are two exceptions to this: A goes to A^* radially from the origin and E goes to E^* radially. The vectors OA and OE are singled out by the fact that the transformation leaves their directions unaltered. Such vectors are called *eigen-vectors*.*

This geometrical description of an eigenvector is easily translated into algebra. The vector having the same direction as v but k times the length is kv. So if $v^* = kv$ for some k, this means that v^* and v have the same direction. This is exemplified in Figure 50 where $A^* = 2A$ and $E^* = 3E$.

A custom has grown up of using the Greek letter lambda, λ, for the number we called k. Thus v is an eigenvector if $v^* = \lambda v$. The number λ is called the *eigenvalue*. Thus A is an eigenvector with the eigenvalue 2. In more informal language, the vector OA gets

*Sometimes also *proper vectors* or *latent vectors*. The German 'eigen' means 'proper'.

stretched to two times its original length (without change of direction). The vector OE is stretched three times. Thus E is an eigenvector with $\lambda = 3$.

Looking for axes that bring the matrix to diagonal form is the same problem as looking for eigenvectors. This is shown by the following considerations.

Suppose first we have found axes that give diagonal form. If the matrix is $\begin{pmatrix} p & 0 \\ 0 & q \end{pmatrix}$ the equations will be $x^* = px$, $y^* = qy$. The transformation is $xc+yd \to pxc+qyd$. Putting $x = 1$, $y = 0$ we see $c \to pc$; putting $x = 0$, $y = 1$ we see $d \to qd$. So c is an eigenvector with eigenvalue p and d is an eigenvector with eigenvalue q.

Now we use the argument the other way round. Suppose we have found two eigenvectors c and d, pointing in different directions. We want to show that if they are chosen as axes, the matrix of the transformation will appear in diagonal form. Let the eigenvalues be p and q. This means $c \to pc$ and $d \to qd$. Then $xc+yd \to xpc+yqd$. So $x^* = px$ and $y^* = qy$. The matrix is thus $\begin{pmatrix} p & 0 \\ 0 & q \end{pmatrix}$, which is in diagonal form.

Note here that the numbers p, q which occur in the diagonal are the eigenvalues, the numbers that tell us how much the eigenvectors get stretched.

It is something of a mouthful to say that a transformation is specified by a matrix in diagonal form; for shortness, we may speak of a transformation being 'given in diagonal form'. This would apply to the transformation $c \to pc$, $d \to qd$, since, as we have just seen, the matrix corresponding to this is diagonal.

When a transformation is given in diagonal form, it is easy to calculate its powers. Let M denote the transformation $c \to 2c$, $d \to 3d$ considered at the beginning of this section. This operation applied twice would give M^2 ; $c \to 4c$, $d \to 9d$. Applied three times, it gives M^3 ; $c \to 8c$, $d \to 27d$. Applied n times it gives M^n ; $c \to 2^n c$, $d \to 3^n d$, for each time M is applied c gets doubled and d gets multiplied by three.

It is also easy to calculate a polynomial involving M. M^2+M, for example, would be $c \to 6c$, $d \to 12d$ for the 'yield' of M^2+M is found by adding the 'yields' of M^2 and M. Indeed, the

arithmetic conceals the pattern of what is happening. If we wrote M^2+M as $c \to (2^2+2)c$, $d \to (3^2+3)d$ we would be able to see how the pattern of x^2+x runs right through this result.

If instead of particular numbers, such as 2 and 3, we use algebraic symbols, the patterns appear immediately. Let T denote the transformation $c \to pc$, $d \to qd$. Then we find T^2; $c \to p^2c$, $d \to q^2d$ and T^n ; $c \to p^nc$, $d \to q^nd$. The expression T^3+5T^2+7T gives the transformation $c \to (p^3+5p^2+7p)c$, $d \to (q^3+5q^2+7q)d$. Here the pattern of the cubic x^3+5x^2+7x is apparent throughout.

The result clearly does not depend on the particular cubic chosen for this example. There is plainly some very general result involved here, and we would like to sum it up by saying that if $f(x)$ denotes any polynomial expression then $f(T)$ is the transformation $c \to f(p)c$, $d \to f(q)d$. There is, however, one small fly in the ointment. What is to happen if the polynomial involves a constant term? Consider a very simple example, $f(x) = x+1$. What is $f(T)$ to mean? If we simply replace x by T we get $T+1$: here we have a *transformation* T added to the *number* 1, and this does not make sense.

Let us look at what we are hoping to get. We want $f(T)$ to give $c \to f(p)c$, $d \to f(q)d$. This last part is perfectly straightforward, even when $f(x) = x+1$. It is $c \to (p+1)c$, $d = (q+1)d$. We want to interpret this as $T+$'something': for $f(x) = x+1$, and to get $f(T)$ we have to replace x by T and 1 by something as yet undecided. Now T is $c \to pc$, $d \to qd$, and we can see pc and qd in the yields of the transformation $c \to (p+1)c$, $d \to (q+1)d$ we are trying to interpret. What else have we there? If we remove pc and qd we are left with $c \to c$, $d \to d$. This is the identity transformation, I. So I is the 'something' we are looking for. We are pleased by this, for I looks rather like 1. We can save our theorem by a little convention. When we go from $f(x)$ to $f(T)$, it is *agreed* that the number 1 is to be replaced by the transformation I. Thus, for example, if $f(x) = x^2+2x+3$, we agree that $f(T)$ is to mean $T^2+2T+3I$.

On this understanding we can assert our theorem; if T is $c \to pc$, $d \to qd$, then $f(T)$ is $c \to f(p)\,c$, $d \to f(q)\,d$. Note that c and d are eigenvectors of $f(T)$, with the eigenvalues $f(p)$ and $f(q)$ respectively. In the matrix for $f(T)$ we would find $f(p)$ and $f(q)$ in the main diagonal, noughts elsewhere.

This argument has been carried through for transformations in two dimensions, but it works equally well in any number of dimensions. For example, in three dimensions, if T were $c \rightarrow pc$, $d \rightarrow qd$, $e \rightarrow re$, then $f(T)$ would be $c \rightarrow f(p)c$, $d \rightarrow f(q)d$, $e \rightarrow f(r)e$.

SOME EXCEPTIONAL TRANSFORMATIONS

We have seen that when a transformation is given in diagonal form, its geometrical meaning is easily seen and algebraic calculations with it are simple. Clearly then, if we are given any transformation, it is very desirable to find axes that show it in diagonal form. The question arises; can this always be done? The answer is; usually, but not always.

Our discussion of this question will show the value of geometrical considerations. The question could be posed and treated purely in terms of algebra, but the resulting calculations are neither simple nor particularly enlightening.

It is much better to look at the question geometrically. If axes can be found to make the matrix diagonal, these axes, as we have seen, must be eigenvectors. That is to say, the transformation stretches them but leaves them unaltered in direction. Accordingly, if there exists some transformation that cannot be brought to diagonal form, this must be because it is impossible to find two directions unaltered by it.

We have already met (without realizing it) an example of such a transformation. It is specified by the matrix A in question 1 at the end of Chapter Three. Its equations are $x^* = x+y$, $y^* = y$. It is illustrated in Figure 51. Every point moves horizontally. The points D, E, F all move one unit to the right. The points G, H, K move two units to the right. The direction OG swings round to O^*G^*; the direction OH swings to O^*H^*. All points that lie above the axis OBC move to the right. All directions that rise from O swing round towards the east. If we extended Figure 51 to show points below the axis, we would find they all moved left, and the corresponding directions, falling from O, swung west. The points O, B, C on the axis stay fixed. The direction of OB is thus unaltered by the transformation.

The direction east–west is the only one unaltered by the transformation. Every other line swings towards it, as when a pair of shears begins to close.

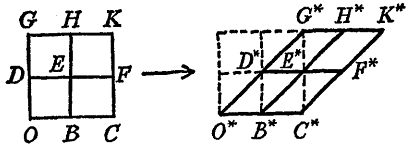


Figure 51

We thus have OB as an eigenvector and as a suitable candidate for one axis, but it is impossible to find an eigenvector lying outside the line OB to serve as the other axis. Accordingly it is impossible to find a pair of axes that show this transformation in diagonal form.

WHEN CAN IT BE DONE?

Matrix theory would be very simple if every transformation could be displayed in diagonal form, but unfortunately this cannot always be done. So the question naturally arises; how can we tell when reduction to diagonal form is possible? It will appear in the next chapter that the equation satisfied by a matrix provides a very simple test.

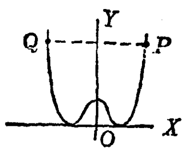
The table shows the equations satisfied by certain matrices from the exercises at the end of Chapter Three. It also shows their eigenvalues and whether or not a change of axes will bring them to diagonal form. This material is rather scanty but by examining it you may be able to make a shrewd guess as to how the equation tells us when diagonal form is possible.

Matrix, X	Equation satisfied	Eigenvalues, λ	Diagonal form exists?
U	$X^2 - I = O$	$+1,\ -1$	Yes
W	$X^2 - 2X = O$	2, 0	Yes
I	$X - I = O$	1	Yes
A	$X^2 - 2X + I = O$	1	No
B	$X^2 - 2X + I = O$	1	No

CHAPTER FIVE

Benefits from Equations

THIS chapter begins with what appears to be a digression but in fact is not. In Figure 52a is shown the graph of $y = x^4 - 2x^2 + 1$.



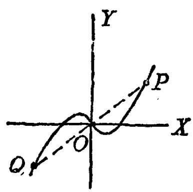


Figure 52

This graph is symmetrical about the line OY; to every point P on the graph corresponds another point Q, its reflection in OY. If P is (x, y) then Q is $(-x, y)$. Figure 52b shows the graph of $y = x^3 - x$. Here to P with coordinates (x, y) corresponds a point Q on the opposite side of the origin with coordinates $(-x, -y)$.

In case (a) we call the graph symmetric, in case (b) antisymmetric. The graph of $y = f(x)$ is symmetric if $f(-x) = f(x)$; it is antisymmetric if $f(-x) = -f(x)$.

These ideas are well known and are much used in the sketching of graphs.

Now of course the majority of graphs are neither symmetric nor antisymmetric. For example, if $g(x) = x^3 + 5x^2 - 7x + 2$, the graph $y = g(x)$ lacks all symmetry. However, there is a theorem that every function can be broken into two parts, one of which is symmetric and the other antisymmetric. The function is the sum of these two parts. For $g(x)$ the symmetric part is $5x^2 + 2$,

the antisymmetric part $x^3 - 7x$. We have simply separated the even powers from the odd ones. For a function such as $g(x)$ the theorem is thus not at all surprising. It is, however, much less evident for graphs such as $y = (x+5)/(2x+3)$ or $y = 10^x$.

Somebody might try to prove this theorem little by little, establishing it first for polynomials, then for fractions, then for infinite series, and so on. This would be entirely the wrong approach. The result does not depend in any way on whether the function involved is simple or complicated. It is just as true for a graph drawn freehand, with no formula at all behind it, as it is for $y = x^2 + x$. The function need not be smooth nor even continuous; it may have sharp bends and jumps. The only restriction is that it must be a function; that is, to each value of x there must correspond just one value of y.

Suppose then we are somehow given the graph $y = G(x)$ and we are seeking to express $G(x)$ as $f(x)+\phi(x)$ with $f(x)$ symmetric and $\phi(x)$ antisymmetric. Let us consider what happens for $x = a$ and $x = -a$, where a stands for any number. Putting $x = a$ we hope to have

$$G(a) = f(a)+\phi(a). \tag{1}$$

Putting $x = -a$, we want to have $G(-a) = f(-a)+\phi(-a)$. Using the facts that $f(-a) = f(a)$ and $\phi(-a) = -\phi(a)$ by the symmetry and antisymmetry of the functions involved, we find

$$G(-a) = f(a)-\phi(a). \tag{2}$$

It is easy to solve the simultaneous equations (1) and (2), in which $G(a)$ and $G(-a)$ are known, and $f(a)$ and $\phi(a)$ are the unknowns we want to determine. We find

$$\left.\begin{aligned} f(a) &= \tfrac{1}{2}\,[G(a)+G(-a)] \\ \phi(a) &= \tfrac{1}{2}\,[G(a)-G(-a)] \end{aligned}\right\} \tag{3}$$

Now here a stands for any number. So, if there are functions f and ϕ that do what is required, they are completely specified by equations (3).

We can check that the functions so specified do in fact meet all requirements. First, using x instead of a, we consider

$f(x) = \frac{1}{2}[G(x)+G(-x)]$. This is symmetrical, for changing x into $-x$ merely alters the order of the terms on the right-hand side. Similarly, we can see that $\phi(x) = \frac{1}{2}[G(x)-G(-x)]$ is an antisymmetric expression, for replacing x by $-x$ changes its sign. Simply adding $f(x)$ to $\phi(x)$ shows their sum to be $G(x)$. So we have succeeded in breaking $G(x)$ into a symmetric and an antisymmetric part.

This solution is shown graphically in Figure 53. The graph of $y = G(-x)$ is got by reversing the graph of $y = G(x)$. The graph

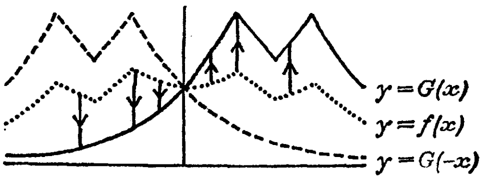


Figure 53

of $f(x)$ then lies midway between these two graphs, for $f(x)$ is the average of $G(x)$ and $G(-x)$. Then $\phi(x)$ has to be added to $f(x)$ so as to give $G(x)$. We do not actually draw the graph of $y = \phi(x)$, but use vertical arrows to show the effect of adding $\phi(x)$. In this particular diagram, we have to raise points of the dotted curve on the right, and lower those on the left, in order to obtain the curve $y = G(x)$. The antisymmetry of $\phi(x)$ appears in the pattern of these arrows.

As already mentioned, we have not had to make any assumption whatever about the nature of the function $G(x)$. The result is entirely due to the *operation* involved, that of changing x into $-x$. So it would seem wise to introduce a symbol, M, for this operation. Thus $MG(x)$ is to mean $G(-x)$. The graphical significance of M may be seen from Figure 53; M turns the graph over, reverses it from left to right.

The operation M is readily carried out when $G(x)$ is specified by a formula. Thus, if $G(x) = x^3+5x^2-7x+2$,

$MG(x) = -x^3+5x^2+7x+2$. But M can also be applied when the function is specified by a table. If we were given:

x	-2	-1	0	1	2
$G(x)$	4	11	7	2	9

we would have:

x	-2	-1	0	1	2
$G(-x)$	9	2	7	11	$4.$

We simply reverse the order of the entries in the second row.

Clearly, if the operation M were applied twice we would be back where we started. So $M^2 = I$.

A symmetric expression, $f(x)$, is one that is unaltered by M. So the condition for symmetry is $Mf(x) = f(x)$.

An antisymmetric expression has its sign changed by the operation M. The condition for this is $Mf(x) = -f(x)$.

In the work above, we began with any $G(x)$ and from it we obtained a symmetric $f(x)$ and an antisymmetric $\phi(x)$.

The symmetric $f(x)$ was given by $\frac{1}{2}[G(x)+G(-x)]$. In terms of M this means $f(x) = \frac{1}{2}[I+M]G(x)$.

How can we see that this $f(x)$ is symmetric? Symmetric means unaltered by M, and this we can demonstrate purely by using algebraic properties of M. For if $f = \frac{1}{2}(I+M)G$, then $Mf = \frac{1}{2}M(I+M)G = \frac{1}{2}(M+M^2)G = \frac{1}{2}(M+I)G$. This last step holds because $M^2 = I$. Thus $Mf = f$.

A similar calculation shows the antisymmetry of ϕ. For $\phi = \frac{1}{2}(I-M)G$. Accordingly $M\phi = \frac{1}{2}M(I-M)G = \frac{1}{2}(M-M^2)G = \frac{1}{2}(M-I)G$. Here too the last step uses $M^2 = I$. Comparing our results for ϕ and $M\phi$ we see $M\phi = -\phi$, which is the test for antisymmetry.

Finally, we have to check $G = f+\phi$. This means that we have to verify $G = \frac{1}{2}(I+M)G+\frac{1}{2}(I-M)G$. The G on the left-hand side of this equation could be written IG. So we could get this equation by starting from the algebraic identity $I = \frac{1}{2}(I+M)+\frac{1}{2}(I-M)$ and then allowing both sides to act on G.

The proofs of the last three paragraphs could be checked by someone who had no idea what M and G stood for, but knew only that $M^2 = I$ and that the usual procedures of algebra were

applicable. The argument above in fact suggests that there must be some very general theorem, running something like this; if $M^2 = I$, any G can be expressed as $f+\phi$, where $Mf = f$ and $M\phi = -\phi$. We cannot leave this statement as it stands, for we must be assuming something about M and G, and what kind of things f and ϕ are.

EXTRACTING THE CONDITIONS

The statement of the theorem mentions $f+\phi$ so at the very least we must be assuming that, in some sense, f and ϕ can be added. By examining the steps of the proof we can see what other assumptions are necessary.

According to the proof, the required f is given by $\frac{1}{2}[I+M]G$. What do we mean by $[I+M]G$? We mean $G+MG$. So MG must be something that can be added to G. Write MG as H for short. We must be able to give meaning to the addition $G+H$. When we have done this, $f = \frac{1}{2}(G+H)$. We must be in some system where multiplication by half is meaningful. The two operations we have considered here are of the type we met in Chapter One with collections of animals – addition, and multiplication by a number. On page 59 we defined a vector space informally as any system in which you could calculate by the rules that apply to collections of animals. Accordingly, it will be sufficient to say that G and H belong to some vector space. This will ensure that $f = \frac{1}{2}(G+H)$ and $\phi = \frac{1}{2}(G-H)$ are meaningful and can be dealt with by the familiar processes of Chapter One. For example, the calculation $f+\phi = G$ is justified.

We have just seen that G and H must belong to the same vector space. Now H is short for MG. So MG belongs to this vector space; that is, M must be an operation mapping this space into itself.

But not any old mapping will do for M. The proof involves Mf, with $f = \frac{1}{2}(G+H)$, and the calculations made implicitly assume results such as $M \,.\, \frac{1}{2}(G+H) = \frac{1}{2}(MG+MH)$.

Any linear transformation has this property. A transformation T is said to be linear if $T(u+v) = Tu+Tv$ and $T(ku) = kTu$. Here u and v may be any vectors and k any number. A 'banking scheme'

automatically has these properties – if you add two investments, you add the yields; if you invest k times as much, the yield will be k times as much. This is why banking schemes were used to introduce the topic of linear transformations.

Any transformation specified by linear equations (or given in matrix form, which means the same thing) is bound to be a linear transformation.

We are now able to state our general theorem, with an explicit description of the situations to which it applies:

If a vector space is given and M is a linear transformation of this space into itself such that $M^2 = I$, then for any vector G we can find vectors f and ϕ such that $G = f+\phi$, where $Mf = f$ and $M\phi = -\phi$.

This sounds very wordy, and to someone unfamiliar with the terminology probably quite terrifying. These words, however, properly understood, are a very concentrated method for reminding us of a rather extensive background. Let us set out their message in full. 'A vector space is given'; this means – do you remember all that stuff in Chapter One about cats and dogs and parallelograms? Do you remember pages 58 and 59 about physicists adding velocities and the pure mathematician's argument that the same formalism would apply to adding quadratics or indeed any functions whatever? If so, you will realize that there are a host of situations to which this theorem could apply; all these situations show certain analogies. 'M is a linear transformation of this space into itself.' In Chapter Three, from page 75 on, we looked at several of these, and we shall in future meet yet more examples of linear transformations. These two statements set the scene: they tell us in what circumstances the theorem will apply. When you remove these stage directions, you are left with the kernel of the theorem, much as we first glimpsed it on page 103.

The theorem just considered is typical of modern mathematics in that *it does not refer to any one particular situation.* It refers to a multitude of situations, all of which have certain aspects in common. This often troubles learners; they are worried by the vagueness, because they cannot clearly imagine all the situations to which the theorem might apply.

To overcome this difficulty it is often helpful to work from both

ends; that is, to consider particular examples and also to consider the logic of the proof.

A simple vector space is the plane. For M we could choose reflection in the horizontal axis, OX, since this makes $M^2 = I$. This example is illustrated in Figure 54. Any point G is chosen. You can think of the vector G either as this point or as the arrow OG, whichever you prefer. H is the reflection of G in OX, so

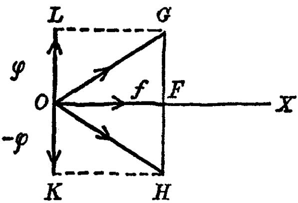


Figure 54

$H = MG$. In the proof of the theorem, we found $f = \frac{1}{2}(G+H)$. Here we recognize the mid-point formula of page 23. The point F is midway between G and H; we can think of f as represented by the point F, or perhaps more conveniently as the arrow OF. Now G should be $f+\phi$, so ϕ is what you must add to f to get G. The vector ϕ could be represented by the arrow FG, or, as in the figure, the arrow OL. When M is applied and the whole figure is reflected in OX, the vector f remains exactly as it was, but ϕ is reversed in direction. So $Mf = f$ and $M\phi = -\phi$ as expected. This example gives a simple but clear illustration of the theorem and helps us to visualize geometrically the algebraic processes used in the proof.

Further examples appear in the next section of this chapter.

As to the other end of the work, while considering particular examples, the learner should always keep coming back to the proof, and seeing that one and the same argument covers all these different cases. In this way any feeling of strangeness tends gradually to disappear.

FURTHER EXAMPLES

We can also obtain geometrical examples from the topic at the beginning of this chapter – symmetric and antisymmetric functions, with M the operation that changes $f(x)$ into $f(-x)$. As we live in only three dimensions, we have to put some restrictions on the functions to be considered if we are to draw pictures.

For a first example, we restrict ourselves to the quadratic expression ax^2+bx+c. The operation M will change this to ax^2-bx+c. If we write a^*, b^*, c^* for the coefficients in this last expression, we have $a^* = a$, $b^* = -b$, $c^* = c$. Fig. 55 shows the effect of this transformation, with a measured east, c north, and b up. We see that the transformation M is simply reflection in the

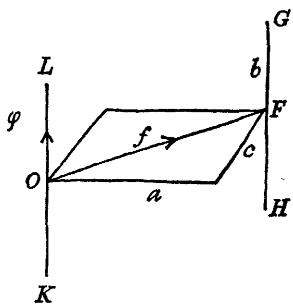


Figure 55

horizontal plane. Any vector OG can be split up into $f = OF$ in the horizontal plane and $\phi = OL$ in the vertical line through O. When M acts, f remains unchanged and ϕ is reversed, so L maps to K.

The horizontal plane and the vertical line through O are called invariant subspaces. Any point, such as L, in the vertical line goes to a point (K) which is also in that line. Any point, such as F, in the horizontal plane goes to a point also in the horizontal plane (actually the same point, F). The plane is its own reflection; the

vertical line KOL is its own reflection. This is why we call them invariant (i.e. unchanging); the operation M leaves them where they were.

These invariant subspaces have a very useful property. If we know what happens in them, we know what happens to any point of space. For we have seen that any point G can be expressed as $f+\phi$, with f lying in the plane and ϕ lying in the vertical line. Knowing what M does to f and ϕ, we can deduce what M does to G, for $MG = Mf+M\phi$. In our figure, $G = F+L$. As $F \to F$ and $L \to K$, we must have $F+L \to F+K$, and $F+K$ is indeed H, the reflection of G.

Note that the invariant spaces here are a plane and a line, both of which are *linear* spaces. This is not peculiar to our present problem; it happens with all linear transformations.

Note also that every vector in the horizontal plane is an eigenvector with $\lambda = +1$, and every vector in the vertical line is an eigenvector with $\lambda = -1$.

Exercise

If instead of quadratics we consider expressions ax^3+bx^2+cx what is the geometrical meaning of M? What are the invariant subspaces, and which values of λ go with the vectors in them? (For answer see page 223.)

If we consider cubic expressions ax^3+bx^2+cx+d, the transformation will be given by $a^* = -a$, $b^* = b$, $c^* = -c$, $d^* = d$. The condition for M to leave a vector unaltered is $a = 0$, $c = 0$. All vectors of the form $(0, b, 0, d)$ form an invariant subspace, which in fact is a plane, since every vector in it is of the form $bu+dv$, where $u = (0, 1, 0, 0)$ and $v = (0, 0, 0, 1)$. Every vector of this plane is an eigenvector with $\lambda = +1$. Similarly, the plane containing all points of the form $(a, 0, c, 0)$ is an invariant subspace; any vector in it is an eigenvector with $\lambda = -1$. Any vector in the whole space of four dimensions can be expressed as $f+\phi$, with f in the first plane and ϕ in the second. We cannot properly visualize space of four dimensions, but it is clear that we have here an analogy with the situation in three dimensions already considered.

The transformation M just considered, with $a^* = -a$, $b^* = b$, $c^* = -c$, $d^* = d$, is already in diagonal form. We do not really

need our theorem about $G = f+\phi$ to study it. However, the theory would be helpful if we were considering broken line graphs such as G in Figure 56. The operation M reverses this graph (compare page 101) and has the equations $a^* = d, b^* = c, c^* = b, d^* = a$. We know that it must be possible to express G as $f+\phi$, where, as

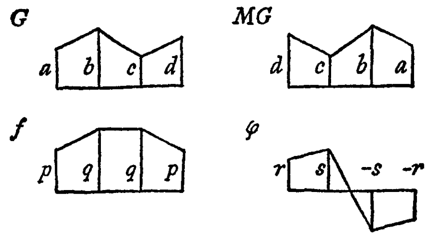


Figure 56

shown in the figure, f is a symmetric graph with entries p, q, q, p and ϕ is an antisymmetric graph with entries $r, s, -s, -r$. This means $a = p+r$, $b = q+s$, $c = q-s$, $d = p-r$. With the new variables p, q, r, s, the transformation reduces to diagonal form, $p^* = p, q^* = q, r^* = -r, s^* = -s$.

The graph G in Figure 56 is determined by the four numbers a, b, c, d. The totality of such graphs form a vector space of four dimensions. If we had taken more points and joined them by straight lines, we could have obtained a vector space with a larger number of dimensions. If we took very many points indeed, with very short pieces of line between them, we would obtain a broken-line graph very closely resembling a continuous curve. There is a suggestion here that the graphs of continuous functions may form a vector space with an infinity of dimensions. We are returning here to an idea we met in Chapter Two, that there are vector spaces, each member of which is a function. You may have noted also that this chapter began with a consideration of the symmetry and antisymmetry of functions. When we analysed the result this led to, we obtained a theorem which began, 'If a vector space is

given . . .'. The idea of a vector space composed of functions is probably not an easy one. It is interesting to see how the study of functions keeps giving us hints of the need for this idea. In Chapter Seven we shall pass from hints to explicit assertion, and try to nail this idea down formally.

GENERALIZING THE EQUATION

So far in this chapter we have been obsessed by the single equation $M^2 = I$. But, as we have seen, there are many other equations that a transformation can satisfy. Can we extract from our particular case principles that will be of general value?

In our work above, we began with *any* vector G, and from it we manufactured* an eigenvector f, given by $\frac{1}{2}(M+I)G$. Why does this particular expression $\frac{1}{2}(M+I)$ turn up here? Since M satisfies the equation $M^2 - I = O$, it is not hard to guess that $M+I$ occurs because it is a factor of $M^2 - I$. How does this help us to show that f is an eigenvector? The condition for an eigenvector (with $\lambda = +1$) is $Mf = f$, which can be written $(M-I)f = 0$. This condition contains the other factor of $M^2 - I$. If we substitute $f = \frac{1}{2}(M+I)G$ in the condition $(M-I)f = O$, the two factors get together to produce $M^2 - I$, which is O. The actual calculation runs as follows. $(M-I)f = (M-I).\frac{1}{2}(M+I)G = \frac{1}{2}(M^2-I)G = O.G = O$.

An idea emerges. Suppose we are dealing with a transformation T that satisfies the equation $(T-I)(T-2I)(T-3I) = O$. We are looking for a vector u that will satisfy $(T-I)u = O$. We can get one by putting $u = (T-2I)(T-3I)G$, where G is any vector whatever. For then $(T-I)u = (T-I)(T-2I)(T-3I)G = O.G = O$. In the same way, if we want a vector v satisfying $(T-2I)v = O$, we choose $v = (T-I)(T-3I)G$. We put all the factors there, except the one that is already present in the given condition.

In our work with $M^2 - I$ we did not choose $(M+I)G$ for f, but rather $\frac{1}{2}(M+I)G$. The extra factor $\frac{1}{2}$ does not spoil the argument. In fact we can multiply our expression by any constant we like

* The vector O satisfies the equation $Mf = \lambda f$ but is not counted as an eigenvector. It can happen that $\frac{1}{2}(M+I)G$ gives O, in which case we have an exception to the statement above, and to similar statements made later on about the manufacture of eigenvectors.

(avoiding 0 of course). So, for the transformation T considered above, it would be quite in order to take $u = a(T-2I)(T-3I)G$ and $v = b(T-I)(T-3I)G$, with any numbers a and b. To complete the list, we want a vector w satisfying $(T-3I)w = O$, and we can choose $w = c(T-I)(T-2I)G$.

In the work with M, an essential point was that G could be broken up into f and ϕ, that is, we had $G = f+\phi$. To keep the analogy, here we shall want to have $G = u+v+w$. This means $G = a(T-2I)(T-3I)G + b(T-I)(T-3I)G + c(T-I)(T-2I)G$. Now this is to happen for every G. This means that the complicated operation on the right-hand side, applied to any vector G, simply leaves it unaltered; that is, it is the identity operator, I. Accordingly we want – if it can be done – to find numbers a, b, c such that:

$$I = a(T-2I)(T-3I)+b(T-I)(T-3I)+c(T-I)(T-2I).$$

This is a problem in traditional algebra. We put it in more familiar form. Can we find numbers a, b, c so that equation (4) below is an identity?

$$1 = a(x-2)(x-3)+b(x-1)(x-3)+c(x-1)(x-2) \qquad (4)$$

This is a standard piece of elementary algebra. It can be solved by multiplying out and solving simultaneous equations or, much more neatly, by taking in turn $x = 1$, $x = 2$, $x = 3$. These substitutions show $a = \frac{1}{2}$, $b = -1$, $c = \frac{1}{2}$ to be the only possibility, and these values do in fact give a solution.

Accordingly, if we choose these values for a, b, c, the equations laid down earlier will enable us to express any vector G in the form $u+v+w$, where $(T-I)u = 0$, $(T-2I)v = 0$, $(T-3I)w = 0$. These last three equations are equivalent to $Tu = u$, $Tv = 2v$, $Tw = 3w$; that is, the transformation T stretches u, v, and w, without any change of direction, one, two, and three times respectively.

Now of course different vectors G will lead to different vectors u, v, w, just as, in Figure 55, different points G can lead to different points F. In fact, in that figure, if we imagine G to vary, the possible positions of F will fill the horizontal plane. In the same way, it can be shown in our present example, that the possible positions of u fill a linear subspace. The possible positions of v fill another linear subspace, and those of w a third. Everything works out in essentially the same way as it did for $M^2-I = 0$.

If we divide equation (4) by $(x-1)(x-2)(x-3)$ and insert the values $a = \frac{1}{2}$, $b = -1$, $c = \frac{1}{2}$ we arrive at the equation:

$$\frac{1}{(x-1)(x-2)(x-3)} = \frac{\frac{1}{2}}{x-1} - \frac{1}{x-2} + \frac{\frac{1}{2}}{x-3}.$$

Anyone who has gone a certain distance with calculus will recognize this as the dissection into partial fractions that would be used to integrate the expression on the left-hand.

The work with M made use of the identity $1 = \frac{1}{2}(1+x) + \frac{1}{2}(1-x)$. If we divide this by x^2-1, we reach the partial fraction result:

$$\frac{1}{x^2-1} = \frac{\frac{1}{2}}{x-1} - \frac{\frac{1}{2}}{x+1}.$$

Someone familiar with partial fractions may find this connexion helpful for remembering the procedure, which may be stated as follows: we have a transformation (or a matrix) M that satisfies an equation $F(M) = 0$. Express $1/F(x)$ in partial fractions. Multiply the resulting equation by $F(x)$. This gives an equation of the form '$1 =$ a certain expression'. Replace x in that expression by M, and apply the result to any vector G.

This procedure will break down in certain cases, namely, when repeated factors occur in $F(x)$. We have seen that the matrix A, mentioned at the end of Chapter Three, cannot be reduced to diagonal form. The simplest equation A satisfies is $A^2-2A+1 = 0$, which corresponds to $F(x) = x^2-2x+1$. Now $x^2-2x+1 = (x-1)^2$. We cannot express $1/(x-1)^2$ in partial fractions in any way that will make it simpler than it now is. Our procedure, in fact, fails when the equation $F(x) = 0$ has repeated roots.

Can we conclude that when this happens the matrix cannot be put into diagonal form? Not yet; we have only shown that our procedure fails – but maybe some other procedure would do the job? We can in fact dispose of this objection. There is a theorem that, if a matrix is expressible in diagonal form, then it satisfies an equation with each root occurring only once.

The proof is not difficult. On page 96 we saw that, if T could be put in diagonal form with p, q, r in the main diagonal, then $F(T)$ would also appear in diagonal form with $F(p)$, $F(q)$, $F(r)$ in the diagonal. Similar results hold in any number of dimensions.

Suppose then, for example, that we are working in six dimensions, so that T is given by a matrix with six rows and six columns. Suppose this matrix is in diagonal form, the diagonal containing the entries 1, 1, 1, 2, 3, 3. Then, by the theorem quoted in the last paragraph, $F(T)$ will be in diagonal form, with diagonal entries $F(1)$, $F(1)$, $F(1)$, $F(2)$, $F(3)$, $F(3)$. If we can choose the function F so as to make all six entries nought, we shall have $F(T) = 0$: for $F(T)$ is a matrix in diagonal form, and all the entries not in the diagonal are nought anyway. But to make the six entries nought, we only have to ensure $F(1) = 0$, $F(2) = 0$, $F(3) = 0$. This can be done by choosing $F(x) = (x-1)(x-2)(x-3)$. With this choice of $F(x)$ no factor is repeated; there is no repeated root.

A formal mathematical proof of this result would simply be a little essay showing that what was done in this particular example could always be done.

Here we have a very satisfactory thing, a necessary and sufficient condition. If a transformation satisfies an equation with no repeated roots, the transformation can be specified by a matrix in diagonal form; if it does not, it cannot.

FINDING THE EQUATION

To apply our last result, we need to find the simplest equation satisfied by a transformation. This raises the question – does every linear transformation satisfy an equation? – if so, how can we find the simplest equation it satisfies? We have to say 'the simplest equation' for it may satisfy many, just as in traditional algebra the number 3 satisfies the simplest equation $x-3 = 0$, but it also satisfies $(x-3)(x-5) = 0$ and indeed any equation $(x-3)f(x) = 0$, where $f(x)$ is a polynomial.

It is easy to show that every linear transformation must satisfy some equation.* We will establish this for two dimensions; the

*We are still thinking in terms of a finite number of dimensions. This statement would be untrue if spaces of infinite dimension were considered. The operation of differentiation, $D = d/dx$ is linear. It passes the tests for linearity given on page 103, for $D(u+v) = Du+Dv$; $D(ku) = kDu$. But there is no equation $F(D) = 0$. Question for investigation: what equation does D satisfy if we make the space finite, e.g. require the functions differentiated to be at most quadratics?

argument is easily generalized. Questions 18–23 at the end of Chapter Three established that 2×2 matrices constitute a linear space of four dimensions. We can appeal to the ideas of the section 'A Useful Result' on pages 55–7. Let T be any 2×2 matrix, and consider I, T, T^2, T^3, T^4. There are five of these and they lie in a space of four dimensions. Thus they must be linearly dependent, that is, connected by an equation $aI+bT+cT^2+dT^3+eT^4 = 0$, with numbers a, b, c, d, e not all nought. So T must certainly satisfy an equation of the fourth degree at most.

In fact, things are far better than the above argument would indicate. For any 2×2 matrix T we find in fact that I, T, T^2 are already linearly dependent, so T satisfies a quadratic equation.

The calculation runs as follows. Suppose $T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then matrix multiplication gives $T^2 = \begin{pmatrix} a^2+bc & ab+bd \\ ac+cd & bc+d^2 \end{pmatrix}$. If we use P, Q, R, S as explained in question 18 of Chapter Three, we may write:

$$\begin{array}{lllll} I = & P & & + & S \\ T = & aP+ & bQ+ & cR+ & dS \\ T^2 = & (a^2+bc)P+ & (ab+bd)Q+ & (ac+cd)R+ & (bc+d^2)S. \end{array}$$

Can we now find numbers m, n so that $T^2+mT+nI = O$? As in the section 'Testing linear dependence' (page 54) we can find m and n so that things start right – that is to say, so that P and Q occur in $T^2+mT+nI$ with coefficients nought. For this gives two equations for the two unknowns, m and n. We can only hope that the coefficients of R and S will look after themselves – that they will turn out to be nought as well. In fact they do. For m and n we find the values $-(a+d)$ and $ad-bc$ respectively. Thus T satisfies the equation:

$$T^2-(a+d)T+(ad-bc)I = 0. \tag{5}$$

Equation (5) is known as the *characteristic equation* of the matrix T. In deriving it, we have not assumed anything about the numbers a, b, c, d. Our result holds for any 2×2 matrix. However,

it need not be the simplest equation satisfied by T. Thus, if $a = 2, b = 0, c = 0, d = 2$, equation (5) becomes $T^2 - 4T + 4I = O$, corresponding to $(x-2)^2 = 0$, an equation with a repeated root. One must be careful not to leap to the conclusion that T cannot be put in diagonal form – for in fact this matrix is already diagonal. T satisfies the simpler equation $T - 2I = O$, corresponding to $x - 2 = 0$, which is free from repeated roots.

TRANSFORMATIONS AND MATRICES

A transformation is closely related to the matrix that specifies it, so closely indeed that many reputable mathematicians use the same symbol to denote both: they speak of the transformation T and of the matrix T. It is important, however, to realize the distinction between these two things. As we have seen, one and the same transformation can be represented by many different matrices, depending on the system of axes used. If some statement is made about a matrix, in general we have no reason to suppose that this statement will remain true if the axes are changed. A statement about a matrix is, so to speak, subjective; it depends on our point of view (our axes). But a statement about a transformation is objective; it corresponds to what on page 114 we call a *geometrical fact*. For example, suppose some transformation T satisfies $T^2 = I$. The transformation, carried out twice, brings everything back to where it started. You cannot get away from this by changing axes. Whatever axes we may choose, we shall find this T represented by a matrix whose square is I. The same argument applies to an equation such as (5). Suppose some transformation satisfies $T^2 = T + I$. Both sides of this equation have geometrical meanings, independent of axes. T^2 simply means that T is applied twice. $T + I$ takes a little longer to describe. For any point P, let $TP = Q$. Then $(T+I)P = TP + IP = Q + P = R$ say, where R is the fourth corner of the parallelogram formed by O, P, and Q. (As was pointed out in the footnote on page 31, the origin O, representing *nothing*, is a geometrical fact.) So T^2 must send P to R, a point specified by a geometrical construction. An argument of this kind applies to any equation satisfied by a transformation T.

Accordingly, equation (5) has an objective meaning. In whatever

axes we may represent T, we shall always arrive at the same characteristic equation. The coefficients in this equation are bound to have some important meaning. You will probably recognize the coefficient of I as the determinant, $ad-bc$, of the matrix for T. The expression $a+d$ which occurs as the coefficient of $-T$ is known as the *character* of T. It is the sum of the elements in the main diagonal, and plays an important role in the beautiful theory of group characters.

A quick rule for obtaining the characteristic equation of T is to write down the determinant of $T-\lambda I$ and equate it to nought. This gives:

$$\begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = 0. \tag{6}$$

If equation (6) is multiplied out, it will be seen to correspond to (5), of course with T replaced by λ and the matrix I by the number 1.

This recipe works equally well in any number of dimensions. In many textbooks the characteristic equation is introduced by means of this determinant. The textbook of course will have to prove that T satisfies the equation so obtained. This result is known as the Cayley–Hamilton theorem. Cayley himself said this was so obvious that he would merely state the result and not bother to give a proof. We do not know exactly what argument was in his mind; mathematicians since Cayley have felt that this theorem did have to be proved, and with some care.

CHAPTER SIX

Towards Applications

THE central theme of Chapter Four was that we could see the effect of a transformation most easily if we knew it in diagonal form. Chapter Five considered when, and how, we could get it into that form. In this chapter we ask what use can be made of the results in those chapters.

Often in nature we can start a process which then develops according to its own laws. It is so when we detonate an explosive, or throw a ball. If we are skilful, we can throw a ball with chosen direction and speed. Once its flight starts, we have no more control over it. The laws of the moving ball are expressed by its *equations of motion*. When we throw it, we choose the *initial conditions*. A *solution* would be a formula showing how the ball would move if started off in any manner whatever.

As a simple analogy, we may consider a Fibonacci sequence. This is a sequence of numbers, in which each number is required to be the sum of the two numbers that precede it. If we call the sequence $a_0, a_1, a_2, a_3, \ldots$ the requirement is written $a_n = a_{n-1}+a_{n-2}$. This, of course, does not help us to choose the first two numbers. Choosing the first two numbers is like starting the flight of the ball. If we choose – as is usually done – the numbers 0, 1, the rest of the series is automatically generated: 0, 1, 1, 2, 3, 5, 8, 13, 21, . . . A solution would be a formula giving the general term a_n for an arbitrary choice of the first two numbers a_0, a_1.

A linear transformation is concealed in the rule for forming the Fibonacci series. The sequence actually written above ends with the two numbers 13, 21. These are all we need to know if we wish to continue the series. The next number must be $13+21 = 34$. If we write 34, the series now ends with 21, 34, and these two numbers are now all we need to know if we wish to continue yet further. The 'genetic code' is thus carried by a pair of numbers, and they generate the next pair: $(13, 21) \to (21, 34)$. Let us denote

the initial pair (0, 1) by (x_0, y_0), the next pair (1, 1) by (x_1, y_1), and so on. We want equations to show how any pair (x_n, y_n) generates the next pair (x_{n+1}, y_{n+1}). In the example (13, 21) → (21, 34) we notice how 21, the second number of the input, reappears as the first number of the output. Quite generally $x_{n+1} = y_n$. The second number of the output, 34, arose as 13+21. Generally $y_{n+1} = x_n + y_n$. Thus $(x_n, y_n) \rightarrow (y_n, x_n + y_n)$. We pass from any pair (x, y) to the next pair (x^*, y^*) by the transformation T; $x^* = y$, $y^* = x+y$.

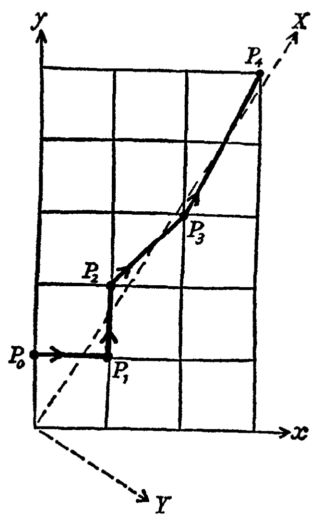


Figure 57

As the series develops, this transformation is applied again and again: (0, 1) → (1, 1) → (1, 2) → (2, 3) → (3, 5) → .... Starting with (x_0, y_0) we have to apply the transformation n times to reach (x_n, y_n). So $(x_n, y_n) = T^n (x_0, y_0)$.

We can plot the pairs (0, 1), (1, 1), (1, 2), (2, 3), etc., on graph paper in the usual way. The resulting points are shown in Figure 57. The chain made by these points looks irregular and uninstruc-

tive. Chapter Four assures us that it will become more comprehensible if we can express T in diagonal form. Now T has (in the language of Chapter Five) $a = 0$, $b = c = d = 1$. Its characteristic equation is thus $T^2 - T - I = O$. The corresponding equation in elementary algebra has roots $\frac{1}{2}(1 \pm \sqrt{5})$, or approximately 1·62 and −0·62. These are not arithmetically convenient; the important thing is that they are distinct – no repeated roots. This means that there are two eigenvectors; their directions can be calculated and are shown by the dotted lines in Figure 57. With the dotted lines as axes the transformation must appear in diagonal form; it is in fact $X^* = 1{\cdot}62X$, $Y^* = -0{\cdot}62Y$. Thus every time T is applied, the X coordinate gets multiplied by 1·62, the Y coordinate by −0·62. See how this shows in Figure 57. The X coordinate grows steadily. But 0·62, the number associated with Y, is less than 1, so the Y value shrinks at each step. Indeed, if a few more points were shown, the Y coordinates would become so small that the points would seem to lie on the X axis. There is also a minus sign involved. At each step Y changes sign. This also appears in the figure. The points skip from one side of the dotted X-axis to the other.

Detailed calculation will show that the dotted X-axis has the equation $y = 1{\cdot}62x$ in the original system. The points (x_n, y_n) get closer and closer to this line. This means that the ratio y_n/x_n approaches 1·62, more strictly $\frac{1}{2}(1 + \sqrt{5})$, as n grows large. That is, the ratios given by successive terms of the Fibonacci sequence, namely, 1/1, 2/1, 3/2, 5/3, 8/5, 13/8, 21/13 . . . approach $\frac{1}{2}(1 + \sqrt{5})$. This is well known to students of classical algebra, but is certainly not obvious to anyone seeing the Fibonacci series for the first time. It is interesting that our theory of linear transformations leads us straight to this result.

THE FIBONACCI PROBLEM

The general problem of the Fibonacci sequence is the following; suppose any two numbers are chosen for a_0, a_1, what will be the formula for the general term a_n? Now that we know the diagonal form of T it is not hard to answer this question. We will sketch the procedure without giving all the details. In Figure 57 we shall now have a chain of points, not starting at $x = 0$, $y = 1$ but at

$x = a_0$, $y = a_1$. This starting point could be specified by means of the dotted axes; let (X_0, Y_0) denote its coordinates in that system. Each time T is applied, the X coordinate gets multiplied by 1·62, the Y coordinate by $-0{\cdot}62$. After n applications, we shall reach point (X_n, Y_n) where $X_n = (1{\cdot}62)^n X_0$, $Y_n = (-0{\cdot}62)^n Y_0$. This tells us the position of the point P_n in Figure 57, but specified in the dotted axes system. It is a routine job to find its coordinates (x_n, y_n) in the original system. Now x_n, y_n stand for the pair of numbers a_n, a_{n+1} in the Fibonacci series; $x_n = a_n$, $y_n = a_{n+1}$. So it will be sufficient to find x_n; this will give us the general term, a_n. If we took the trouble to calculate y_n, this ought to confirm our formula, but would not yield any new information.

The change of axes leads to a result of the form $x_n = A\,(1{\cdot}62)^n + B\,(-0{\cdot}62)^n$. The numbers A and B do not depend on n, but of course they do depend on the choice of a_0 and a_1.

Accordingly, the formula for any Fibonacci sequence must be of the type $a_n = A(1{\cdot}62)^n + B(-0{\cdot}62)^n$. It is perhaps surprising that the formula for the sequence 0, 1, 1, 2, 3, 5, 8 . . . , consisting entirely of whole numbers, should depend so much on the irrational number $\sqrt{5}$.

DIFFERENCE EQUATIONS

We have not chosen the Fibonacci sequence merely for its interest. The Fibonacci rule $a_n = a_{n-1} + a_{n-2}$ is an example of a *difference equation*. Difference equations are important in themselves and are also intimately related to *differential equations* which play a great role in applications of mathematics. The result we found for the Fibonacci sequence is typical of what happens in a wide region.

Many books give rules for solving difference equations. Learners do not like these rules, which seem completely arbitrary. Suppose we want to solve the difference equation $a_n = 5a_{n-1} - 6a_{n-2}$. *Rule One* says, 'Try to find a solution which is a geometrical progression; that is, try $a_n = r^n$.' If we try this on our difference equation, we are led to the equation $r^2 = 5r - 6$, which has the solutions $r = 2$ and $r = 3$. Accordingly, we have the two special solutions, $a_n = 2^n$ and $a_n = 3^n$. *Rule Two* now tells us to combine these special solutions by writing $a_n = A.2^n + B.3^n$. This is the

general solution. In fairness to the textbooks, it must be admitted that they prove the correctness of Rule Two. It is Rule One that puzzles learners; why should we look for solutions of the form r^n?

One thing should be made perfectly clear. Rules One and Two do give, for practical purposes, the simplest and quickest way of arriving at the solution. Our discussion here is not aimed at finding a better algorithm. Its aim is rather to provide a rational background, to show how one could arrive at such rules. For nothing destroys mathematical ability more quickly than the habit of accepting procedures without asking *why* this procedure is followed, *how* we might have arrived at it for ourselves.

If we studied the equation $a_n = 5a_{n-1} - 6a_{n-2}$ in the way we have already studied the Fibonacci sequence, we would find that this equation was associated with the transformation T; $x^* = y$, $y^* = -6x+5y$. T has eigenvalues 2 and 3, corresponding to the eigenvectors (1, 2) and (1, 3). If we based our axes on these eigenvectors, the transformation would take the form $X^* = 2X$, $Y^* = 3Y$. Applying this transformation n times would multiply the X coordinate by 2^n and the Y coordinate by 3^n. Finally, we would reach the formula $a_n = x_n = A.2^n + B.3^n$ on returning to our original system of axes.

After working out one or two cases in this way, and seeing why such a formula arose, learners could use Rules One and Two when they actually needed to solve a difference equation.

What we have done leads us to expect solutions of the type $a_n = Ap^n + Bq^n$, where p and q are eigenvalues, and this is normally what happens. However, there are exceptional cases and our approach leads us to expect these, *for a transformation cannot always be put in diagonal form*. Consider the sequence of numbers 0, 1, 2, 3, 4, 5, . . . They satisfy the equation $a_n = 2a_{n-1} - a_{n-2}$ but it is very hard to believe (and is indeed untrue) that they can be given by a formula $a_n = Ap^n + Bq^n$. The transformation T here is $x^* = y$, $y^* = -x+2y$, which satisfies $(T-I)^2 = O$, but no simpler equation. Thus the equation has repeated roots and so T cannot be expressed in diagonal form. We will not here pursue the details of what happens in this special case.

FROM DIFFERENCE TO DIFFERENTIAL EQUATIONS

At the beginning of this chapter we took as a simple illustration of mathematical physics the flight of a ball. Now this motion is (or seems to us to be) continuous; we see the stone moving continuously and not by jumps or jerks. If we recorded the motion with a ciné-camera we would be introducing jumps. One frame would show the stone in one place, the next in another. There would be no indication of how it moved in between. If we numbered the frames 0, 1, 2, . . . we would have a sequence of points $P_0, P_1, P_2 \ldots$ where P_0 shows the position of the stone in frame 0, P_1 in frame 1, and so on.

We frequently find ourselves in the position where we can cope with some continuous process only by replacing it by a sequence. For example, any real positive number x has a logarithm $\log_{10}x$. But we cannot conceivably produce a table in which the logarithms of *all* numbers between 1 and 10 are shown. Every table of logarithms contains a finite sequence of numbers. It may be a small table giving the logarithms of 1, 1·01, 1·02 . . . or an ambitious table giving the logarithms of 1, 1·0000001, 1·0000002, and so on. But either way, it gives only a sequence.

All classical physics is based on the idea of continuity, and its characteristic weapon is the *infinitesimal* calculus. Equations of motion are almost invariably expressed by differential equations. These make statements about such things as velocity, acceleration, steepness, and curvature. Everything depends on the idea of a limit – what happens when some quantity approaches zero.

But suppose we do not go to the limit. Suppose we make a high-speed film with pictures taken every thousandth of a second. Surely from such a film we could get a very good idea of how an object was moving. Surely we could make a very good estimate of its velocity and acceleration? If we followed this idea through, instead of continuous functions we should have sequences, and instead of differential equations we should have *difference equations*, the topic with which this chapter has so far been concerned.

Pure mathematicians react unfavourably to this idea. Suppose, for example, such a film showed an object in exactly the same place

in each picture. We would tend to assume it was at rest. But suppose it moved away after each picture had been taken and raced back into position just before the next shot? Logically, such a thing is perfectly possible. Our information does not prove the object to be at rest; it does however rather suggest it. We could satisfy the mathematician by explicitly stating a number of assumptions that would rule out such awkward possibilities.

ERSATZ CALCULUS

We proceed to develop this idea – to see what kind of formulas we would get for velocity, acceleration, and so forth in a world that moved by jumps. This procedure has two purposes. It allows us to explain (in a rather crude way) the meaning of differential equations and partial differential equations to a person completely ignorant of calculus. It is also relevant to the way in which electronic computers deal with calculus situations, for a digital electronic computer has this in common with a table of logarithms – it can deal only with sequences; it has no equipment for dealing with continuous changes; it moves by jerks.

Our discussion may throw some light on calculus, but it should be made perfectly clear that it does not replace calculus. In fact, paradoxically, the ideas of genuine calculus have to be appealed to frequently if one is not to get incorrect results from an electronic computer. The basic ideas of calculus remain something that should be taught to as many people as possible while they are still as young as possible.

It is fairly clear how we would estimate velocity from a film. If a ball appeared on one picture at a height of 7 feet and on the next picture, taken 0·001 of a second later, at a height of 7·003 feet, we would estimate that it was rising at 3 feet a second. We are dividing distance gone by time taken. In the same way, if our camera recorded events at time intervals of h and showed the successive positions of an object at distances $a, b, c, d, e \ldots$, the velocity record of this object would be estimated as

$$\frac{b-a}{h}, \frac{c-b}{h}, \frac{d-c}{h}, \frac{e-d}{h} \quad \ldots.$$

We could go on to estimate acceleration. Acceleration stands in the same relation to velocity that velocity does to distance, so we merely repeat the process already used – subtract each entry from the next and divide by h. This gives the acceleration record as

$$\frac{c-2b+a}{h^2}, \frac{d-2c+b}{h^2}, \frac{e-2d+c}{h^2}, \ldots.$$

In calculus, if the distance s is given in terms of time t by $s = f(t)$, the velocity is denoted by ds/dt or $f'(t)$ and the acceleration by d^2s/dt^2 or $f''(t)$. Calculus applies not only to movements but also to graphs. If we have the graph $y = f(x)$ then dy/dx or $f'(x)$ measures the steepness of the graph, and d^2y/dx^2 or $f''(x)$ tells us how the curve is bending. Where $f''(x)$ is positive, the curve resembles a bowl, where $f''(x)$ is negative, an arch.

We would expect to find some similar graphical interpretation for estimated velocity and acceleration found above, and in fact

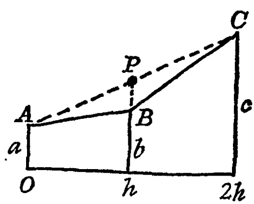


Figure 58

we can. In Figure 58 $(b-a)/h$ and $(c-b)/h$ measure the steepness of the lines AB and BC. The expression $(c-2b+a)/h^2$, the estimate of $f''(x)$, is interpreted geometrically as follows. In Figure 58, P is the midpoint of AC. As A is $(0, a)$ and C is $(2h, c)$, P must be $\frac{1}{2}A + \frac{1}{2}C$ (see p. 23). So P is $(h, \frac{1}{2}a + \frac{1}{2}c)$. As B is (h, b), the height of P above B is $\frac{1}{2}a + \frac{1}{2}c - b$, that is, $\frac{1}{2}(c-2b+a)$. So $(c-2b+a)/h^2$ is $2/h^2$ times the height BP. When it is positive, P is above B, as in Figure 58. When it is negative, P is below B, as in Figure 59. If it should happen to be nought, P and B coincide; this means that the points A, B, C are in line.

Our approach leads to one difficulty that does not arise in genuine calculus. The quantity $(b-a)/h$ represents the steepness at every point of the line AB, so we do not know whether we should regard it as an estimate of $f'(0)$, the steepness at A, or $f'(h)$, the steepness at B. Sometimes one is used, sometimes the other, whichever seems more convenient. The decision is rather arbitrary. We are less arbitrary in regard to $(c-2b+a)/h^2$. We have seen that this is proportional to the length BP, so it is specially related to B, the

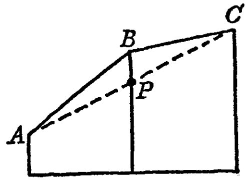


Figure 59

middle of the three points involved. Accordingly we regard it as an estimate of f'' at B, rather than at A or C.

Figure 58 could be imagined as representing part of a stretched string, such as we considered in Chapter Four. The tension of a string tries to straighten the string, so B would feel itself pulled towards the point P, which is in line with A and C. It is not surprising that detailed calculations show, for a string with unit tension, the force pulling B upwards to be $(c-2b+a)/h$. This is our estimate of $hf''(x)$. As the particles are spaced at distance h, we are led to guess that the upward force per unit length of string might be $f''(x)$, and this is indeed a correct result in the genuine calculus treatment of a string *slightly* displaced from equilibrium.

If, as is usually done in the technical literature, we use a_0, a_1, a_2, a_3, a_4 . . . instead of a, b, c, d, e . . . , our estimates appear as

$$\left.\begin{aligned} f'(nh) &\sim (a_{n+1}-a_n)/h \\ f''(nh) &\sim (a_{n+1}-2a_n+a_{n-1})/h^2 \end{aligned}\right\}. \qquad (1)$$

The sign $\sim$ indicates 'is estimated as' or 'approximately equals'. The first equation here treats $(b-a)/h$ as an estimate of $f'(x)$ at A

in Figure 58. When we decide to regard it as an estimate of $f'(x)$ at B, we have to use the following equations:

$$\left.\begin{aligned} f'(nh) &\sim (a_n - a_{n-1})/h \\ f''(nh) &\sim (a_{n+1} - 2a_n + a_{n-1})/h^2 \end{aligned}\right\}. \tag{2}$$

By using these approximations we turn any problem in calculus into a problem in algebra. A differential equation, connecting $f(x)$, $f'(x)$, and $f''(x)$, is thus replaced by an ordinary algebraic equation connecting a_{n-1}, a_n and a_{n+1}. A problem in calculus is thus reduced to a problem in arithmetic; the solution is found by calculations essentially similar to those by which we found the first few terms of the Fibonacci sequence, 0, 1, 1, 2, 3, 5 . . . .

For example, the differential equation $f'' + 3{\cdot}7f' + 3f = 0$ could arise in a system where a mass is connected to a spring, but its motion is opposed by some treacly liquid – an arrangement like a shock absorber. Suppose we take $h = 0{\cdot}1$, corresponding to a film made with ten pictures a second. Using equations (2) above, the differential equation will be replaced by:

$$100\,(a_{n+1} - 2a_n + a_{n-1}) + 37(a_n - a_{n-1}) + 3a_n = 0$$

since $1/h^2 = 100$ and $1/h = 10$. This equation boils down to:

$$a_{n+1} = 1{\cdot}6a_n - 0{\cdot}63a_{n-1}.$$

We can, if we like, handle this difference equation in a primitive way, using arithmetic only. We would assume values for a_0 and a_1; then a_2 would be found by substituting in $a_2 = 1{\cdot}6a_1 - 0{\cdot}63a_0$, then a_3 would be found from a_2 and a_1, and so on. Alternatively, we can use the rule for solving difference equations given earlier in this chapter. We would find $a_n = A(0{\cdot}9)^n + B(0{\cdot}7)^n$. If we started the process off by taking $a_0 = 0$, $a_1 = 0{\cdot}2$, this would mean $A = 1$, $B = -1$ and so $a_n = (0{\cdot}9)^n - (0{\cdot}7)^n$. Either by this method, or by the primitive approach, we would arrive at the sequence 0, 0·2, 0·32, 0·386, 0·416, 0·422, 0·413, 0·395, ... After this the numbers would gradually return to nought.

In Figure 60, the little circles represent the numbers just calculated. The curve represents the exact theoretical solution of the original differential equation. Even with the very crude approximations we have used, we have at any rate obtained a correct general

picture of the way in which the mass would move. By using $h = 0{\cdot}01$ or $h = 0{\cdot}001$ we could obtain a better approximation. It can be proved (for this differential equation) that, by taking h small enough, we can achieve any required degree of accuracy.

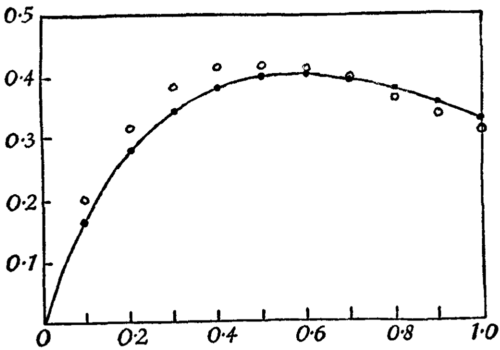


Figure 60

A NOTE ON THE DIFFERENTIAL EQUATION

The exact solution of the differential equation, $f'' + 3{\cdot}7f' + 3f = 0$, considered above, is $f(x) = Ae^{-1{\cdot}2x} + Be^{-2{\cdot}5x}$. This solution is usually obtained by a rule much like that for solving difference equations, except that Rule One now reads, 'Look for a solution of the form $f(x) = e^{mx}$'. Learners raise exactly the same objection – 'Why should we do this?' Our earlier explanation is sufficient to show why it is natural to do so. First of all, the differential equation is a limiting case of the type of difference equation we studied earlier; it is reasonable to *guess* (we are not speaking of proof) that methods that worked earlier in this chapter may also work now. Second, what we are doing now is exactly the same as what we did then. If we write $e^m = r$, then $e^{mx} = r^x$, and r^x corresponds to the r^n we used earlier. It would be perfectly possible and correct to solve differential equations by trying solutions of the form r^x. The form e^{mx} is completely equivalent to this, but is more convenient for use in the formulas of calculus.

PARTIAL DIFFERENTIATION

Partial differentiation is a curiously neglected subject. Many people who know the basic ideas of calculus are completely unaware of it. Yet it is widely useful and, for someone who has met calculus, does not involve any essentially new idea or formula. We will now try to convey some idea of what partial differentiation is about, without even drawing on any of the results of calculus.

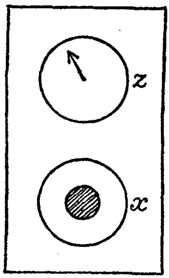


Figure 61

Figure 61 illustrates a simple piece of apparatus. At the bottom a dial can be turned; the reading of the dial is called x. Some kind of measuring instrument is visible at the top; its reading is called z. The meter and the dial are linked by some mechanical or electrical system, so that the dial setting fixes the pointer reading. Thus $z = f(x)$. In ordinary calculus, dz/dx or $f'(x)$ denotes the rate at which the pointer reading would change if the dial were rotated at unit speed, i.e. in such a way that the dial reading, x, increases by 1 every second. Now consider the instrument shown in Figure 62. There is still a meter, with reading z, which now depends on the setting of two dials, with readings x and y. We write $z = f(x, y)$. How are we now to talk about the rate at which z changes? There are many different ways in which the two dials might be whirling around. We pick out two standard situations.

In the first, the y-dial is held fixed and the x-dial is turned at unit speed. The rate at which the z reading then increases is denoted by $\partial z/\partial x$, and is called the *partial* derivative of z with respect to x. In the second standard situation, the x-dial is held fixed and the y-dial rotates at unit speed. The resulting rate of change of z is called $\partial z/\partial y$.

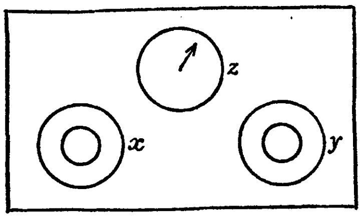


Figure 62

A change of image is now possible. Suppose (x, y) to be co-ordinates on a map, and $z = f(x, y)$ to represent the height of the ground above sea level at the point (x, y). In the first standard situation, y is held fixed and x increases at unit rate. This corresponds to walking east at unit speed. So $\partial z/\partial x$ measures the rate at which your height above sea level increases as you walk due east at unit speed. This is the same thing as the steepness, or gradient, of the ground in the easterly direction (see Figure 63). In the same way, the second standard situation corresponds to walking north at unit speed, and $\partial z/\partial y$ measures the gradient of the ground in the northerly direction.

It is just as easy to estimate $\partial z/\partial x$ and $\partial z/\partial y$ as it was earlier to estimate $f'(x)$.

If we were surveying a hilly landscape it would be impossible for us to record the height of the land at *every* point. We might cover our map with a grid, such as that shown in Figure 64, and measure the heights only at the points where two lines crossed.

If the grid had a sufficiently small mesh and the landscape was not unduly jagged, this would give us a good impression of the lie of

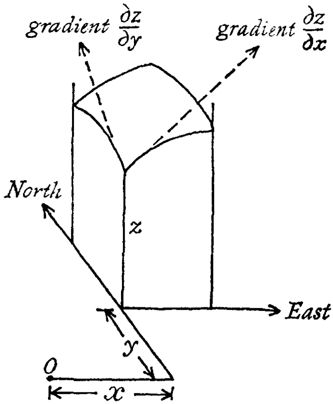


Figure 63

the land. We now agree to write a for the height of the land at A, b for the height at B, and so on. We suppose the grid to consist of squares* of side h.

If we walk from A to B, we cover a horizontal distance h and

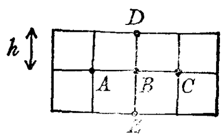


Figure 64

we rise a height $b-a$. Accordingly we estimate the gradient as $(b-a)/h$. As we are walking in an easterly direction, this is an estimate for $\partial z/\partial x$. In the same way $(b-e)/h$ measures the gradient of EB and gives us an estimate for $\partial z/\partial y$.

*For some problems squares are unsatisfactory and rectangles have to be used.

We may repeat the process; $\partial z/\partial x$ tells us how fast z changes as we move east. We may ask in turn; how fast does $\partial z/\partial x$ change as we move east? The gradient of AB is $(b-a)/h$. If we move over a distance h to the east, the gradient of BC is $(c-b)/h$. The difference between these is $(c-2b+a)/h$. Since this difference is caused by a shift h to the east, to find the rate of change we must divide by h. We obtain $(c-2b+a)/h^2$. This is a familiar expression; we met it earlier as an estimate for $f''(x)$. In fact we are dealing with the same situation. If we made a vertical section of the countryside along the line ABC, we would arrive at a figure much like Figure 58 on page 123. We usually write $\partial^2 z/\partial x^2$ for the quantity we have just estimated. This resembles d^2z/dx^2 used in beginning calculus. Replacing d by ∂ indicates that another variable, y, is involved but that we are holding it fixed and allowing only x to vary.

In the same way we use the symbol $\partial^2 z/\partial y^2$ to indicate the corresponding quantity for a section of the countryside by a vertical plane in a northerly direction. Our crude estimate for $\partial^2 z/\partial y^2$ is $(d-2b+e)/h^2$.

In both these estimates, the letter b occurs in the middle. We regard them as estimating the values of $\partial^2 z/\partial x^2$ and $\partial^2 z/\partial y^2$ at the point B.

AN ELECTRICAL PROBLEM

It is rather remarkable that, with the rather scanty considerations just given, we can make out a plausible case for the equation satisfied by the flow of electricity in a continuous copper sheet. It is

$$\frac{\partial^2 V}{\partial x^2}+\frac{\partial^2 V}{\partial y^2}=0.$$

This is a famous equation, occurring in a dozen or more branches of science, and associated with the name of Laplace. The expression on the left-hand side of the equation is often referred to as the Laplacian of V, and sometimes abbreviated to ΔV or $\nabla^2 V$.

We suppose we have a flat piece of thin copper sheet. It is uniform; it has the same thickness and electrical resistance throughout. Various batteries are connected to points on its

boundary, and as a result currents flow in the sheet. We want to investigate the distribution of these currents.

We begin by simplifying the problem. If we look at a handkerchief, it at first appears as a continuous surface. Closer examination shows that it consists of woven threads. It seems likely that if the material of a handkerchief were somehow changed into copper, the electrical properties of the resulting object, woven out of copper wires, would closely resemble those of a continuous sheet of copper. Accordingly instead of a continuous sheet, we consider a grid of fine copper wires, much like the grid of lines shown in Figure 64. We hope this will not alter the problem too much.

Next we have to consider how electricity flows in a net made of wires. Two laws operate here, both fairly simple. The first law states that electricity flows *through* the material, much as water flows through pipes. That is to say, all the current that flows into

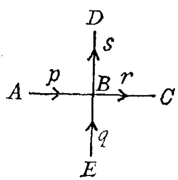


Figure 65

any point has to flow out again. For example, in Figure 65 we see currents p and q flowing into the point B, and currents r and s flowing out. The total out must equal the total in, so $r+s = p+q$.

The second law is Ohm's law. At every point there is a number V, called the potential. Potential is something like height for gravity or temperature for heat. Water flows from high places to low places; heat from places where the temperature is high to places where it is low; electricity from points with high potential to points with low potential. We write a, b, c, d, e for the potentials at A, B, C, D, E. If current flows from A to B, as in Figure 65, the potential a must be larger than the potential b. The potential

drop from A to B is $a-b$ volts. According to Ohm's law, the current flowing from A to B is proportional to this potential drop. For simplicity we will assume that the current from A to B is not merely proportional but actually equal to $a-b$. Thus $p = a-b$. Similarly, we shall have $q = e-b$, $r = b-c$, $s = b-d$. On substituting in the equation $p+q = r+s$ given by the first law we find $a-2b+e = 2b-c-d$.

We can handle this equation in various ways. If we solve for b we find $b = \frac{1}{4}(a+c+d+e)$. This is an interesting result; it says that the potential at B is the average of the potentials at the four neighbouring points A, C, D, E.

Again, the equation can be written $(c-2b+a)+(d-2b+e) = 0$. This is highly reminiscent of our estimates for $\partial^2 z/\partial x^2$ and $\partial^2 z/\partial y^2$. If we divide the equation by h^2 we find:

$$\frac{c-2b+a}{h^2} + \frac{d-2b+e}{h^2} = 0. \tag{3}$$

This suggests:

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0. \tag{4}$$

This is Laplace's Equation.

STRETCHED NETS AND SOAP BUBBLES

It was mentioned earlier that Laplace's Equation had many different applications. One of these involves something rather

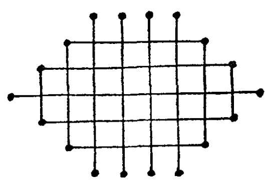


Figure 66

like a warped tennis racket. Figure 66 represents a piece of net, tightly stretched. Imagine it first of all lying on a table. Then

suppose the boundary points to be raised various small distances from the table and clamped in position. The weight of the net is negligible. We now regard Figure 65 as representing part of the net. The quantities a, b, c, . . . now represent the heights of A, B, C, . . . above the table. By the remark made on page 124, $(c-2b+a)/h^2$ is proportional to the upward force exerted by the string ABC on the point B. (It is in fact $1/h$ times the force, if the net has unit tension.) In the same way $(d-2b+e)/h^2$ is proportional to the upward force on B exerted by the string EBD. Now if the net has settled down to its equilibrium position, the total upward force on the point B must be zero. This is exactly what equation (3) states. So equation (3) is the condition for the net to be in equilibrium. If we imagine a succession of nets with finer and finer meshes these will approach in the limit a continuous membrane, like a drumhead or a soap bubble. We conjecture that equation (4) will give the equilibrium condition for these.

There is a theorem that, if V satisfies Laplace's Equation in a certain region, then V cannot have a maximum or a minimum anywhere inside this region. We can see the reasonableness of this theorem, and indeed in two ways. We saw that the value of V at B was the average of its values at the four neighbouring points A, C, D, E; this is clearly incompatible with a maximum existing at B. We can also see the result physically, by considering our stretched net. If B is higher than all its neighbours, then all the strings BA, BC, BD, BE are pulling B downwards; the point B could not possibly remain at rest in this situation. Actually, when a net is in equilibrium the two 'bendings' must be in opposite directions. By this I mean that if the string ABC is tending to lift B (like the string in Figure 58) then the string EBD must be tending to lower B (like the string in Figure 59). If ABC is like a bowl, EBD must be like an arch, and vice versa.

We can use the considerations above in various ways. The problem of the stretched net or membrane is easy to visualize; it may help us to see the meaning of Laplace's Equation, and indeed to guess properties of that equation. For example, if we clamp all the points on the boundary of the net in position, our physical experience tells us that this will determine the position of the whole net. If we warp the frame of a tennis racket in a

particular way, this determines how the strings will come to rest. We should expect there to be some mathematical theorem corresponding to this, according to which the values of V on the boundary fix the values of V throughout the interior, if V is a solution of Laplace's Equation. And indeed there is a theorem, involving certain reasonable conditions on V, which states precisely this.

Again, since the same equation has so many different applications, we can use one branch of science to help another. We can solve a problem in electricity by observing the shape of a soap bubble.

We can also use these ideas for computing numerical solutions. The algebraic equation (3) is related to the partial differential equation (4) in the same way that a difference equation is related to a differential equation.

STABILITY

There is a trap, of which everyone who does numerical work, whether with a pencil, desk calculator, or electronic computer, should be aware.

We can illustrate this trap by means of the very simple differential equation $y' = -y$. Now of course no one would dream of solving this numerically. The solution is $y = Ae^{-x}$ and tables of e^{-x} are readily available. We use it merely to illustrate a danger that lurks everywhere.

We begin by considering a correct treatment. We use $h = 0{\cdot}1$ and formula (1) of page 124, so that y is replaced by a_n and y' by $(a_{n+1} - a_n)/(0{\cdot}1)$. This gives us the difference equation $(a_{n+1} - a_n)/(0{\cdot}1) = -a_n$, which simplifies to $a_{n+1} = 0{\cdot}9\, a_n$. Thus each entry will be nine tenths of the previous one; the entries will decrease steadily. If we are told that the initial value of y is 1,000, we put $a_0 = 1{,}000$. Working to the nearest whole number we would obtain the sequence 1,000, 900, 810, 729, 656, ... These numbers of course differ appreciably from the values of the exact solution $y = 1{,}000e^{-x}$, as a result of the very crude approximation we have used. It might occur to us to try to improve things in the following way. In Figure 67 the true value of y' at the point B is the gradient of the tangent BT. We have replaced it by

$(a_{n+1} - a_n)/(0{\cdot}1)$ which corresponds to the gradient of BC. It looks as though the gradient of AC would give a much better estimate of the gradient of BT. And indeed it would. Thus, if B

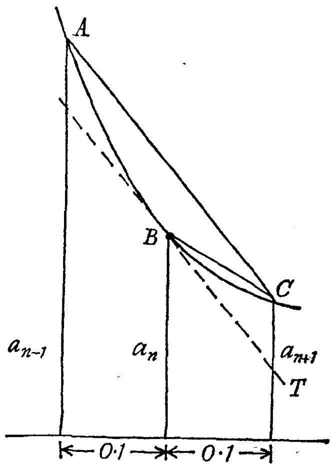


Figure 67

happened to be the point with $x = 1$, the true value of y' would be $-0{\cdot}3679$ approximately. The gradient of BC would be $-0{\cdot}3501$, with an error of $0{\cdot}0178$; the gradient of AC would be $-0{\cdot}3685$, with an error of $0{\cdot}0006$ only.

It looks therefore as if we would obtain better results if we replaced y' by the gradient of AC, which is $(a_{n+1} - a_{n-1})/(0{\cdot}2)$. This would lead to the equation $(a_{n+1} - a_{n-1})/(0{\cdot}2) = -a_n$, which simplifies to $a_{n+1} = a_{n-1} - 0{\cdot}2\ a_n$.

The table on page 136 shows the values of a_n as calculated from this difference equation, starting with $a_0 = 1{,}000$ and $a_1 = 900$. The table also shows the values of the true solution, $1{,}000\ e^{-x}$, and the errors involved in replacing these by a_n.

x	n	a_n	$1{,}000\ e^{-x}$	Error
0	0	1,000	1,000	0
0·1	1	900	905	−5
0·2	2	820	819	+1
0·3	3	736	741	−5
0·4	4	673	670	+3
0·5	5	601	607	−6
0·6	6	553	549	+4
0·7	7	490	497	−7
0·8	8	455	449	+6
0·9	9	399	407	−8
1·0	10	375	368	+7
1·1	11	324	333	−9
1·2	12	311	301	+10
1·3	13	262	273	−11
1·4	14	259	247	+12
1·5	15	210	223	−13
1·6	16	217	202	+15
1·7	17	167	183	−16
1·8	18	184	165	+19
1·9	19	130	150	−20
2·0	20	158	135	+23

This table shows a marked wobble. The numbers a_n are alternately too large and too small, and *the error steadily increases.* The values ought to decrease steadily, as they do in the column of exact values, but the wobble affects the estimates a_n so badly that a_{16} is actually bigger than a_{15}, while a_{18} and a_{20} are considerably bigger than their predecessors.

If we had been relying on this calculation for our ideas about the solution of $y' = -y$, we would have been seriously misled. The wobble is entirely the result of our method of calculation; it has no basis in reality.

What has gone wrong? Where has the wobble come from? If we solve our difference equation $a_{n+1} = a_{n-1} - 0{\cdot}2\ a_n$ by the method explained earlier, we find $a_n = K\,(0{\cdot}905)^n + M(-1{\cdot}105)^n$. Now this solution *can* give an excellent approximation to 1,000 e^{-x}. If we take $K = 1{,}000$ and $M = 0$, we get the sequence 1,000, 905, 819, 741, . . . which, to the degree of accuracy we have been using, is in exact accord with the true values. It is the best result

we have had yet. This reflects the fact that, in Figure 67, the chord AC gives a much better approximation than BC to the direction of the tangent at B. We ought to benefit from the extra care in our approximation.

The trouble is that we only get this good result if we are able to make M *exactly* nought. Our table that had the wobble in it was based on the initial values $a_0 = 1{,}000$, $a_1 = 900$. These values would result if we put $K = 997\frac{1}{2}$, $M = 2\frac{1}{2}$. This may seem a small change from $K = 1{,}000$, $M = 0$, for an alteration of $2\frac{1}{2}$ is not very big when 1,000 is involved. The trouble arises from the fact that, in the formula for a_n, the coefficient of M is $(-1{\cdot}105)^n$. We are not particularly concerned with the effect of the minus sign here. What is important is the size of this term; its magnitude is $M(1{\cdot}105)^n$. Now this is the formula for something growing at a rate rather more than ten per cent. However small it may be at the beginning, it will end by being very large.

The position thus is that all is well so long as $M = 0$, but if an error is made that has the effect of making M differ from nought by the tiniest amount, then this error will grow and grow until it swamps everything else.

Now it is in the nature of most calculations that some error is inevitable. We can perhaps make exact calculations in the Fibonacci sequence, where all the numbers 0, 1, 1, 2, 3, 5 . . . are whole numbers. But most problems involve fractions; some involve irrational numbers. No machine can cope with an infinite decimal. If we are working, say, to four places of decimals, and we want to use the fraction $\frac{1}{3}$, we alter it to 0·3333. If in the course of some work we need to multiply 0·2345 by 0·1111 the true answer is 0·02605295; this gets cut down to 0·0261. An electronic computer may work to ten places of decimals; even so, it is forced to introduce some errors. If in arriving at an answer it carries out millions of operations, unavoidably it introduces millions of errors.

All computation involves a complication not found in theoretical mathematics. In theory, all quantities have exact values; these grow and develop like a person following a course on an open plain. But computing is like navigating in a forest; we have a course that we try to follow, but we may find a tree in the way, so we dodge to the right or the left. The question is whether the

place where we emerge will be determined by our navigating instructions, or by the thousands of small detours we have made on the way.

An equation such as $a_{n+1} = a_{n-1} - 0{\cdot}2\, a_n$, in which any error made is predestined to grow indefinitely, is called *unstable*.

THE NEED FOR CAUTION

Anyone applying the methods of this chapter should be aware of two possible sources of error.

Our basic idea was to replace a process of continuous change by a process involving jerks. We study what happens with the jerky process. We suppose the jerks to become ever smaller and more rapid. We hope that our solution of the jerky problem will then approach the solution of the continuous problem. Often it does, but not always. There are cases, not at all suspicious in their appearance, in which the jerky solution approaches something entirely unlike the desired solution of the original continuous problem.

The other danger is that already pointed out in the previous section, the possibility of instability.

A result obtained with the help of a computer is only to be taken seriously if the magnitudes of the errors that may arise from these two causes have been carefully estimated.

There is a story (the origin of which I cannot trace) according to which a man processed some meteorological data and forecast a typhoon. No typhoon came. The supposed typhoon was entirely due to an inappropriate numerical procedure.

CHAPTER SEVEN

Towards Systematic Classification

IN previous chapters we have freely used algebraic operations in dealing with vectors, matrices, linear transformations. We have formed powers T^n; we have added and multiplied matrices; we have found equations satisfied by matrices. In all of this we have used the familiar routines of elementary algebra and been led to useful results. Certain anomalies, however, have been present. For example, in elementary algebra the equation $x^2 = 1$ has two and only two solutions, $x = +1$ and $x = -1$. The corresponding matrix equation $M^2 = I$ admittedly has the solutions $+I$ and $-I$, but it also has a host of others. In the first exercise at the end of Chapter Three, we saw that $U = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ makes $U^2 = I$. If $K = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ we also have $K^2 = I$. In fact there are infinitely many 2×2 matrices that satisfy $M^2 = I$.

This leaves us feeling rather insecure. Evidently there are many situations in which elementary algebra leads us to correct conclusions about matrices, but there are also occasions where it is completely misleading. When we have finished such a calculation we do not know whether to believe the result or not. The matter clearly needs tidying up.

Why does a quadratic equation behave so differently when matrices are involved? Let us consider how, in elementary algebra, we would prove that $x^2-1 = 0$ has only the two solutions $+1$, -1. We might argue as follows:

(1) $x^2-1 = (x-1)(x+1)$.

(2) $\therefore (x-1)(x+1) = 0$ if x is a solution of $x^2-1 = 0$.

(3) A product is nought only when at least one of its factors is nought.

(4) So $x-1$ or $x+1$ must be nought.

(5) So x must be 1 or -1. Q.E.D.

Now consider $U = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. This satisfies $U^2 - I = O$ but U is neither I nor $-I$. At which step would the matrix argument, corresponding to the proof above, fail?

Step (2) checks all right. For $U - I = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$ and $U + I = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. Multiplying these two matrices together, we do verify $(U-I)(U+I) = O$.

The argument must fail by step (4), since neither of the matrices $U - I$ and $U + I$ found above is in fact O. And indeed it is principle (3) that fails. The product $(U-I)(U+I)$ is O, but neither factor is O.

Accordingly, in matrix work we cannot rely on any result in elementary algebra, the proof of which depends on principle (3).

But this is not the only property that fails. Consider the following argument, where $U = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ as before and $K = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

(6) $K^2 = I$ and $U^2 = I$ so $K^2 - U^2 = O$.
(7) $(K-U)(K+U) = K^2 - U^2$.
(8) $\therefore (K-U)(K+U) = O$.
(9) So $K - U$ or $K + U = O$.
(10) So $K = U$ or $K = -U$.

The conclusion is certainly false. The step from (8) to (9) involves principle (3) and we are tempted to believe this is where the argument goes wrong. But the argument goes astray before this; equation (8) is already incorrect. For $K - U = \begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix}$ and $K + U = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, so $(K-U)(K+U) = \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}$ which is not O.

Now statement (6) is all right, so the culprit must be (7). And in fact, equation (7) is incorrect. The right-hand side is O. The left-hand side $(K-U)(K+U)$, as we have just seen, is different from O. So with matrices the familiar factoring of the difference of two squares is no longer permissible – at any rate not as a general rule.

So we look back a stage further. How is $(a-b)(a+b) = a^2 - b^2$ proved in traditional algebra? Simply by multiplying out the left-hand side. We try this with K and U. $(K-U)(K+U) =$

$K(K+U)-U(K+U) = K^2+KU-UK-U^2$. In elementary algebra, we would cancel KU and UK, and have the desired result. But with matrices this is not permissible. By matrix multiplication we find $KU = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and $UK = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. These are not equal.

Note that we could complete the above calculation exactly as in elementary algebra, if we happened to be dealing with two matrices A and B for which $AB = BA$, for instance $A = \begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}$. For *these* matrices, we should be justified in asserting $(A-B)\,(A+B) = A^2-B^2$.

When $AB = BA$, we say that A and B *commute*.

If M is any matrix whatever,* M commutes with the identity matrix I, and also with any power, M^n, of itself. We have $MI = IM = M$ and $M.M^n = M^n.M = M^{n+1}$. This has the welcome result that polynomials in a single matrix are multiplied in exactly the same way as polynomials in elementary algebra, so that all the familiar formulas survive. For example, we are quite safe in asserting $(M-I)\,(M+I) = M^2-I$. This is why the statement (1) in the first proof of this chapter carried over quite successfully for the matrix U.

Various steps, permissible in elementary algebra, cannot be applied to matrices. For example, given $ax = ay$ with $a \neq 0$, we are accustomed to conclude $x = y$. But with matrices this may not work. For example, if $A = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$, $X = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix}$, $Y = \begin{pmatrix} 4 & 5 \\ 6 & 7 \end{pmatrix}$, both AX and AY equal $\begin{pmatrix} 2 & 2 \\ -2 & 2 \end{pmatrix}$, and $A \neq O$, but of course X is not equal to Y. This is connected with the fact that a matrix product can be zero without either factor being zero. For $AX = AY$ is equivalent to $A(Y-X) = O$. Earlier in this chapter we met two non-zero matrices, $U-I$ and $U+I$ whose product was O. The example just given was in fact constructed by choosing A, X, Y so that $A = U-I$ and $Y-X = U+I$.

* M is assumed square, i.e. with as many rows as columns, so that M^n is defined.

When $AB = O$ without either A or B being O, we call A and B *divisors of zero*.

THE CLASSIFICATION OF MATHEMATICAL SYSTEMS

We see that we have to exercise caution in working with matrices because some matrices (but by no means all) are divisors of zero, and because often (but not always) a pair of matrices fail to commute.

Very often we are concerned only with matrices of a particular type. For example, we might restrict ourselves to matrices of the type $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$. If P and Q are of this type, then certainly $PQ = QP$ and you never find $PQ = 0$, unless of course P or Q happens to be 0. Thus this system is commutative and contains no divisors of zero. This system in fact is that of the complex numbers, and it has *all* the algebraic properties possessed by the real numbers. This makes it very easy to work with. Of course we do not often have the luck to find such a system. Usually we have to be content with one that has some of the familiar properties. A notable example is provided by the 4×4 matrices of the type

$$\begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix}.$$

This system is not commutative, but that is its only shortcoming. It has no divisors of zero. It is the system of quaternions. The matrix shown represents the quaternion $a+ib+jc+kd$.

A rotation about the origin (in Euclidean geometry) is represented by the matrix $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, where the numbers a, b are real and satisfy $a^2+b^2 = 1$. If two such matrices are multiplied together, the result is another matrix of the same kind. The matrices commute. Division is always possible; given A and B in this system we can always find C so that $C = A/B$; by this we understand $CB = A$. However, if we are to remain within the system, we must exclude addition. The sum of two of our matrices

will not, in general, represent a rotation. (If we tried to introduce addition, we should find ourselves back at the matrices for complex numbers.) So here we have a system in which there is one basic operation only, multiplication. Multiplication and its opposite, division, behave in the manner we are used to.

A rather larger collection consists of all matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $ad-bc = 1$. Here again we can multiply and divide, but addition and subtraction are barred, since the sum of two such matrices will not as a rule satisfy the condition $ad-bc = 1$. Multiplication is not commutative, so we have to recognize two kinds of division. Given A and B, we can find C so that $CB = A$, and D so that $BD = A$, but in general C and D will be distinct.

These four examples are drawn from a multitude of mathematical systems that have been studied. It is clearly desirable to have some way of classifying mathematical systems, and names have been devised for this purpose. A system which, like the complex numbers, has *all* the stock algebraic properties is called a *field*. The finite arithmetics, discussed in Chapter Thirteen of *Prelude to Mathematics*, provide another, very different, example of a field. A system which, like quaternions, meets all the requirements except commutative multiplication is called a *skew field*. Rotations are an example of a *commutative group*, also referred to as *Abelian groups* in honour of the mathematician Abel. Matrices with $ad-bc = 1$ form simply a *group*.

To earn any of these titles, a system must pass certain tests. If we know it has passed these, we can deduce certain properties it must have. For instance, we can prove, once and for all, that in any field a quadratic cannot have more than two solutions. Knowing some system to be a field,we know something about the behaviour of quadratics in it. We do not have to keep proving the same result over and over again for one system after another.

One reason for studying matrices is that so many different mathematical systems can be exhibited in the form of matrices. However, we need not be obsessed by matrices. We can consider any system whatever and seek its place in our classification. For example, we may consider what pigeon-holes are appropriate for such systems as the integers, the rational numbers, the real num-

bers, the even integers, the operations involving differentiation or integration, and so forth.

We now list some of the tests that can be applied. For each system, we would record which of the following statements are true for it.

(1) For any two elements a, b of the system, there is an element, also belonging to the system, defined as $a+b$.

(2) Addition is commutative; $a+b = b+a$ always.

(3) Addition is associative; $a+(b+c) = (a+b)+c$ always.

(4) The system contains an element 0 such that $a+0 = a$ for every a in the system.

(5) Subtraction is defined; that is, for every a, b, the equation $a+x = b$ has exactly one solution.

(6) For any a, b in the system there is an element, also in the system, defined as ab.

(7) Multiplication is commutative; $ab = ba$ always.

(8) Multiplication is associative; $a(bc) = (ab)c$ always.

(9) The system contains an element I such that $aI = Ia = a$ for every a in the system.

(10) Division is defined; that is, for every a, b the equations $ax = b$ and $ya = b$ each have exactly one solution, provided $a \neq 0$.

(11) Distributive laws: (i) $a(b+c) = ab+ac$, (ii) $(b+c)a = ba+ca$ always.

(12) Freedom from divisors of zero; $ab = 0$ can only happen if $a = 0$ or $b = 0$.

A system that passes all twelve tests qualifies as a field. The tests are not entirely independent; a system that passes tests (1) to (11) is certain to pass (12). The rational numbers, the real numbers, the complex numbers are examples of fields.

A system that passes all tests, except perhaps (7), qualifies as a skew field; e.g. quaternions.

A system that passes (6), (8), (9), and (10) qualifies as a group; e.g. rotations about the origin; all 2×2 matrices with $ad-bc = 1$; the positive real numbers; the real number system with 0 removed; the pair of numbers $1, -1$; the numbers 10^n, where n is any integer (positive, negative, or zero). All of these are sometimes called *multiplicative groups*. This is rather a bad name, for it describes the

way the group is written, rather than its actual nature. It means that we are calling the operation 'multiplication'. But the formal properties of multiplication are the same as those of addition. Notice how (7) and (8) echo (2) and (3). It is this similarity of addition and multiplication that makes logarithms possible – an invention for converting multiplication into addition.

This similarity means that we have a kind of redundancy of symbols. As a result, a convention has grown up that $+$ is used only for commutative systems. Multiplication may be used either for commutative or non-commutative systems. This convention is the reason why tests (4) and (5) do not run exactly parallel to (9) and (10).

A commutative group, using $+$, is thus a system that passes tests (1) to (5); examples – the integers; all integers exactly divisible by 7; any vector space.

If we prefer the multiplication sign – as we would, for example, with the matrices representing rotations – the tests to be passed are (6) to (10).

RINGS

One very important type has not yet been mentioned. Consider the system of all even numbers, ... -4, $-2, 0, 2, 4, \ldots$. It qualifies as a commutative group using $+$, but this does not do full credit to it, for multiplication, as well as addition, is defined in this system. It does not go to the length of being a field, for it fails (10); division is not possible within it. For example, 6 and 2 both belong to it, but $6 \div 2$ is 3, not an even number, while $2 \div 6$ is even worse, the fraction $\frac{1}{3}$. For systems in which one can add, subtract, and multiply, but not necessarily divide, the name *ring* has been devised. It is a peculiar word for this purpose and I do not know how it came to be chosen. The tests a ring must pass are (1) to (6), (8) and (11). The system of all 2×2 matrices constitutes a ring. So does the system of all polynomials in x. If we have two polynomials, say $x+3$ and x^2+1, we can add, subtract, or multiply these; however, division would lead to $(x+3)/(x^2+1)$ which cannot be expressed as a polynomial. So polynomials form a ring but not a field. The integers form a ring but not a field for very

similar reasons. Another system that is a ring but not a field is that of all continuous functions. If $y = f(x)$ and $y = \phi(x)$ are continuous curves, the same will be true for the graphs of the sum, difference, and product, namely, $y = f(x)+\phi(x)$; $y = f(x)-\phi(x)$; $y = f(x)\,\phi(x)$. But we cannot say the same for the quotient graph, $y = f(x)/\phi(x)$. For example, $f(x) = x$ and $\phi(x) = 2x-1$ are continuous, but the quotient $y = x/(2x-1)$ has a break in its graph at $x = \frac{1}{2}$.

It should be noticed that we apply the tests positively and not negatively. A candidate is not disqualified for passing tests other than the required ones. Thus, for example, every field qualifies as a ring. This is reasonable, for in the theory of rings we prove theorems that are true for any system that passes tests (1) to (6), (8) and (11). Such theorems will naturally be true for systems that pass all these tests and some others as well. For example, in any ring we can prove $(x^3-x)(x^3+x) = x^6-x^2$. This result holds true for any field; for example, it holds for the real numbers. The situation here is like that in geometry, where any theorem in affine geometry is also a theorem for Euclid, but not conversely.

Frequently 'ring' occurs with some qualifying adjective. The even numbers form a *commutative ring*, for they pass test (7) as well as the tests compulsory for a ring. The 2×2 matrices do not qualify as a commutative ring. However, 2×2 matrices contain the matrix I, and so pass test (9); they form a *ring with unit element*. The even numbers do not qualify for this title.

One can play endless games of specifying systems and seeing their places in the classification. For example, we could consider 2×2 matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, where a, b, c, d were required to be even integers. This is a ring, but it has neither of the additional properties just mentioned; it is not commutative and it does not have a unit element.

CALCULATION IN RINGS

The results we can prove for rings naturally have a strong resemblance to the results in elementary algebra. In elementary algebra we meet results such as $(1+x)^4 = 1+4x+6x^2+4x^3+x^4$ and

$(x+y)^2 = x^2+2xy+y^2$; we also have generalizations of these, the binomial theorem for $(1+x)^n$ and $(x+y)^n$. Ring theory leads to precisely the same results, but with a difference of scope. In elementary algebra, the binomial theorem is understood to announce the truth of a certain formula *for any numbers x, y*. Ring theory, on the other hand, announces the truth of this formula *for any system that passes certain tests.*

Since 1 occurs in $(1+x)^n$, the system must contain a unit element. If x and y do not commute, $(x+y)^2$ can be brought to the form $x^2+xy+yx+y^2$, but not to $x^2+2xy+y^2$. (See the answer to exercise 9 at the end of Chapter Three.) So we must demand commutativity. We can now state the binomial theorem in a much wider form; the usual formulas for $(1+x)^n$ and $(x+y)^n$ hold in any commutative ring with unit element. (Actually, the first of these holds even when the ring is not commutative, and the second even when there is no unit element.)

The proof of the binomial theorem in this wider sense would follow exactly the same lines as in traditional algebra. The only difference would be that the calculation would be broken down into a lot of small steps, each of which would be justified by referring to the appropriate test which the system is known to have passed. Incidentally, we would need to use results from the traditional algebra of numbers in this proof, for in effect we are trying to show *how many times* (say) x^3y^{n-3} will appear when we multiply $(x+y)^n$ out completely.

I must say I do not find the details of this proof terribly exciting. The interesting thing is the result – we establish our right to apply the binomial theorem to all kinds of things that are not numbers.

AN APPLICATION

A problem that may arise in actuarial and in some forms of scientific work is that of fitting a simple formula to a set of data. Suppose we are given the following information; for some function $f(0) = 0$, $f(1) = 2$, $f(2) = 10$, $f(3) = 30$, $f(4) = 68$, $f(5) = 130$, $f(6) = 222$. We believe a simple polynomial may fit these data; is this so, and, if so, what is the formula?

It is usual to begin by making a table of differences.

n	0	1	2	3	4	5	6
$f(n)$	0	2	10	30	68	130	222
$\Delta f(n)$	2	8	20	38	62	92	
$\Delta^2 f(n)$	6	12	18	24	30		
$\Delta^3 f(n)$	6	6	6	6			
$\Delta^4 f(n)$	0	0	0				

The numbers in each row show the steps by which the numbers in the row above increase. The appearance of only noughts in the row for $\Delta^4 f(n)$ is an indication that a simple formula is involved. How to find it? Actuarial textbooks give the following method.

First we define an operation E which turns $f(n)$ into $f(n+1)$. Thus $Ef(n) = f(n+1)$. In particular $Ef(0) = f(1)$, $Ef(1) = f(2)$, $Ef(2) = f(3)$. Accordingly $f(0)$ can be changed into $f(3)$ by applying the operation E three times. In symbols, $E^3 f(0) = f(3)$. Quite generally, $E^n f(0) = f(n)$. If we can find a formula for $E^n f(0)$ we have reached our goal.

Now what about the operation Δ? Consider a particular number in the table, say 38, which is $\Delta f(3)$. This number arises because it is the difference, $68-30$, of two numbers in the row above; in fact it is $f(4)-f(3)$. Now $f(4) = Ef(3)$, so our number 38 is $Ef(3)-f(3)$, which we may write $(E-1)f(3)$. So we have $\Delta f(3) = (E-1)f(3)$. This suggests that the operation Δ is the same as the operation $E-1$. There was nothing special about the number 38 we picked in the $\Delta f(n)$ row. Exactly the same argument could be carried through for any other entry in that row. Accordingly we accept as true that Δ and $E-1$ represent the same operation. We thus have $\Delta = E-1$, and so $E = 1+\Delta$.

Now we apply the Binomial Theorem:

$$f(n) = E^n f(0) = (1+\Delta)^n f(0)$$

$$= \left\{1 + n\Delta + \frac{n(n-1)}{1.2}\Delta^2 + \frac{n(n-1)\,(n-2)}{1.2.3}\Delta^3 + \dots\right\} f(0)$$

$$= f(0) + n\Delta f(0) + \frac{n(n-1)}{1.2}\Delta^2 f(0) + \frac{n(n-1)\,(n-2)}{1.2.3}\Delta^3 f(0) + \dots .$$

In our particular example, $\Delta^4 f(0)$ and later terms are all nought,

so we need only the terms actually printed in the last expression.

On substituting from the first column of our table $f(0) = 0$, $\Delta f(0) = 2$, $\Delta^2 f(0) = 6$, $\Delta^3 f(0) = 6$, we find:

$$f(n) = 0+2n+3n(n-1)+n(n-1)(n-2),$$

which simplifies to $f(n) = n^3+n$, the formula behind the data.

I remember meeting this method as a boy when reading for the first examination of the Institute of Actuaries. At that time I still thought of the Binomial Theorem as a result about numbers. This method struck me as an ingenious and daring example of formalism, an example of what liberties you could take in mathematics, and still get away with a correct answer. Actually the method is perfectly rigorous and can be logically justified. We have defined the operator E. We can go on to define any polynomial in E, such as E^2+2E+3. By $(E^2+2E+3)f(n)$ we of course understand $E^2f(n)+2Ef(n)+3f(n)$, which is $f(n+2)+2f(n+1)+3f(n)$. The sum of two such polynomials is easily defined. Multiplication is defined by successive application; $(E+2)(E+3)$ is the operation that first applies $E+3$ and then applies $E+2$ to the result. Having made the meanings of our symbols clear, we have to satisfy ourselves that polynomials in E do constitute a commutative ring with unit element. Once having done this – and it is tedious rather than difficult – we shall be justified in applying the binomial theorem to this system. Since $\Delta = E-1$, the operation Δ belongs to the system, so it is legitimate to apply the binomial formula to $(1+\Delta)^n$.

A similar idea is used in one method for handling differential equations. Differentiation is denoted by D, so $Df(x)$ means $f'(x)$. By $(D^2+2D+3)f(x)$ we understand $f''(x)+2f'(x)+3f(x)$. Multiplication is again defined by successive application; $(D+2)(D+3)$ means that the operation $D+3$ is to be applied, and then $D+2$ applied to the result. It is found that polynomials in D form a commutative ring with unit element. Once again, all the formulas of elementary algebra that involve only addition, subtraction, and multiplication can be used with complete confidence. Some textbooks are unnecessarily coy about the method. They say, 'The D method is a way of guessing the solution of a differential equation; you must always test the correctness of the results it gives.' It then

appears as a strange coincidence that the results invariably are correct.

The situation is essentially altered if we begin to consider operations such as $D+x$, where $(D+x)f(x)$ means $f'(x)+xf(x)$. We still have a perfectly good meaning for, say, $(D-x)(D+x)f(x)$. This indicates that the operations $D+x$ and $D-x$ are successively applied to $f(x)$, giving $(D-x)\,u(x)$ where $u(x) = (D+x)f(x)$. It would, however, be wrong to assume $(D-x)(D+x) = D^2-x^2$. For $u(x) = f'(x)+xf(x)$, so $Du(x) = u'(x) = f''(x)+xf'(x)+f(x)$, whence $(D-x)\,u(x) = u'(x)-xu(x) = f''(x)-x^2f(x)+f(x) = (D^2-x^2+1)f(x)$. This means $(D-x)(D+x) = D^2-x^2+1$, a novel result for anyone accustomed to elementary algebra. The difference is due to the fact that we are no longer dealing with a commutative system, for xD and Dx are not equal. xD means 'differentiate, multiply the result by x', so $xDf(x)$ means $xf'(x)$. On the other hand, Dx means 'multiply by x, differentiate the result', so $Dxf(x) = \{xf(x)\}' = xf'(x)+f(x)$ by the rule for differentiating a product. This last result may be written $\{xD+1\}f(x)$, so $Dx = xD+1$. If we bear this last equation in mind, we can multiply out $(D-x)(D+x)$ by an algebraic process, as follows:

$$\begin{aligned}(D-x)(D+x) &= D(D+x)-x(D+x) \text{ by } 11\text{(ii)}\\ &= D^2+Dx-xD-x^2 \text{ by } 11\text{(i)}\\ &= D^2+1-x^2 \text{ since } Dx-xD = 1.\end{aligned}$$

It is good to get used to calculations of this kind, in which it makes a difference if the order of factors in a product is changed. Property (7), commutative multiplication, $ab = ba$, is one that we very frequently have to do without. For example, the system consisting of all 2×2 matrices has some resemblance to the x, D system; it too is a ring with unit element, but not commutative.

One of the main features of quantum theory is that multiplication is not commutative. It is interesting to note that quantum theory can be presented either in terms of matrices (Heisenberg, 1925) or in terms of differentiation (Schrödinger, 1926).

Exercises

1. Calculate $(D+x)(D-x)$. Is it the same as $(D-x)(D+x)$?
2. Calculate $(D+x)^2-(D^2+2xD+x^2)$.

3. Let $X = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ and $I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

   Calculate X^2 and X^3. Calculate $(I+X)^2$ (1) directly by matrix multiplication, (2) from the usual formula for the square of a sum. Do these results agree? Could one predict, before working them out, whether they would agree or not? Find $(I+X)^{10}$.
4. If A and B belong to a *non-commutative* ring, what would $(A+B)^2$ and $(A+B)^3$ give when fully multiplied out?
5. The 'Last-Digit System' is explained as follows. The only symbols allowed are 0, 1, 2, 3, 4, 5, 6, 7, 8, 9. Addition and multiplication are as in ordinary arithmetic, except that only the last digit of the answer is taken. Thus $3+9 = 2$, since 12 ends in 2, and $4\times 7 = 8$, since 28 ends in 8. Which tests does this system pass? To what type does it belong? Would the binomial theorem hold in this system?
6. The 'Even Last-Digit System'. The rules are the same as in question 5, except that only the even digits 0, 2, 4, 6, 8 are used. The same three queries are to be answered.
7. The 'Odd Last-Digit System'. Again, everything as in question 5, except that this time only the odd digits 1, 3, 5, 7, 9 are permitted.

VECTOR SPACES

So far we have talked about vector spaces with only a very loose explanation of what we meant. We will now discuss precisely worded tests for deciding whether or not any system is a vector space.

Our work with the plane used expressions such as $K+4P$. These involved *points* or *vectors*, denoted by the capital letters K, P. They also involved numbers. We shall use small letters to indicate arbitrary numbers; thus $2K+3P$ is an example of the type $xK+yP$.

Our tests fall into two parts. In the first part numbers are not mentioned at all; we are concerned solely with vector addition. In the second part, the multiplication of vectors by numbers is considered.

The first part is short and simple. It contains only five tests.

(1) For any vectors P, Q a vector called the sum, $P+Q$, is defined.

(2) Addition is commutative, $P+Q = Q+P$.
(3) Addition is associative $P+(Q+R) = (P+Q)+R$.
(4) There is a vector O such that $P+O = P$ for every vector P.
(5) Subtraction is defined. For any two vectors P, Q there exists exactly one vector X such that $P+X = Q$.

It will be agreed that these properties hold for the systems we have so far considered, whether a vector is interpreted as a point in a plane with parallelogram addition or as x cats and y dogs.

These tests, incidentally, correspond exactly to tests (1) to (5) on page 144. They require the vector system to be a commutative group.

To arrive at the second part of the testing scheme, we consider how we make calculations when both vectors and numbers are involved. When we add $2P$ to $3P$ we get $5P$; that is, we simply add the numbers involved. The algebraic specification of the process used here is $aP+bP = (a+b)P$.

Again five times $P+Q$ is $5P+5Q$. The general rule is $a(P+Q) = aP+aQ$.

Also, three times $2P$ is $6P$. We have multiplied the numbers involved. The general rule is $a.(bP) = (ab)P$.

Occasionally we need to use the fact that one times P is P; as an equation $1.P = P$.

Now, of course, none of these things could be done if we had not first laid down what multiplying a vector by a number was to mean.

Accordingly, we have the remaining tests as follows:

(6) For every vector P and every number a, the vector aP is defined.
(7) $aP+bP = (a+b)P$ always.
(8) $a(P+Q) = aP+aQ$ always.
(9) $a.(bP) = (ab)P$ always.
(10) $1.P = P$.

If we know that some system passes all these tests, we can prove that all the steps we normally take in working with vectors are justified if used with that system. So it is possible to obtain a body of method and results that apply to *any* system that passes these ten tests.

Towards the end of Chapter Two we suggested that such systems

as that of all quadratics, that of all expressions $a \sin x + b \cos x$, that of all continuous functions, should be regarded as vector spaces. We are now in a position not merely to suggest but to prove these statements. We shall achieve this if we check that each of the systems mentioned has all the properties (1) to (10).

To do this is easy. If we have two quadratic expressions, their sum is defined and is a quadratic. So quadratics pass test (1). If P denotes the quadratic $px^2 + qx + r$, by aP we understand $apx^2 + aqx + ar$. Thus aP is defined and we have met the requirements of test (6). The remaining tests merely call for the checking of simple identities in elementary algebra.

The system consisting of all expressions $a \sin x + b \cos x$ can be disposed of in much the same way.

To prove that continuous functions constitute a vector space, we begin by observing that if $f(x)$ and $g(x)$ are continuous functions, their sum $f(x) + g(x)$ is also continuous. This seems reasonable; a formal proof is a not very difficult exercise in analysis. The requirements of test (1) are thus met; if the word *vector* is interpreted as meaning 'continuous function', the sum of two vectors is defined, and is a vector. Test (6) requires that the multiplication of a vector (a continuous function) by a number a be defined, and that it be a vector (a continuous function). Now if $f(x)$ is continuous, so is $af(x)$. Accordingly, if P represents $f(x)$, we define aP as $af(x)$, and meet the requirements of test (6). All the other properties now follow. The zero, O, required for test (4) is the function whose graph is the x-axis, i.e. $y = 0$ for all x.

In the discussion of continuous functions, it does not matter whether we require our function to be continuous for every real number x, or whether we require this only for some interval, say for x between 0 and 1. Either way we get a vector space.

The system consisting of all polynomials is a vector space. The sum of two polynomials is always a polynomial, so test (1) is passed; if we multiply a polynomial by any number a we obtain a polynomial. The other properties are easily verified. Writing down a formal proof may give some difficulty; the difficulty is as much one of English composition as of mathematics. We have the general idea, but find difficulty in expressing it.

The power series for e^x has the property that it is convergent for

all values of x, real or complex. Series with this property are said to define *entire functions* or *integral functions*. These series, or the functions they define, form a vector space.

We have just discussed continuous functions $f(x)$. Here we have only one variable x; we are setting out from a space of one dimension. But one dimension does not offer any special advantages. Suppose $V = f(x,y)$ gives (say) electric potential at a point (x,y) in a piece of copper sheet and $V = g(x, y)$ the potential in some other circumstances. Then $V = f(x,y) + g(x,y)$ will also specify a potential. For if $f(x,y)$ and $g(x,y)$ are solutions of Laplace's Equation, so is $f(x,y) + g(x,y)$. So also will $af(x,y)$ be a solution, for any number a. We are thus able to define addition and multiplication by a without going away from possible potentials, so the requirements of tests (1) and (6) are met. The other properties are easily checked. Thus the possible electric potentials in a given flat piece of copper (on the understanding that batteries are connected only at the boundary) form a vector space.

We could get another vector space, analogous to the space of all continuous functions $f(x)$, if we did not require $f(x,y)$ to specify a potential, but accepted *any* continuous function $f(x,y)$ defined for all points of some region in the plane.

Vector spaces are thus not a new thing. Anyone adding polynomials or convergent series or considering electrical potentials is dealing with a vector space. The only new thing is the manner in which we classify our experiences.

CHAPTER EIGHT

On Linearity

LINEAR MAPPINGS

IN Chapter Three we considered various mappings that could be represented by matrices. Some were mappings from one space to another (petrol $\rightarrow$ money), some from a space to itself. But all were linear. As explained on page 103, this means they all have the property that, if $P \rightarrow P^*$ and $Q \rightarrow Q^*$, then $P+Q \rightarrow P^*+Q^*$ and $kP \rightarrow kP^*$ for any number k. This definition of linearity implies that we know what is meant by $P+Q$ and kP; these expressions are meaningful if P and Q lie in some vector space. Similarly, P^*+Q^* and kP^* will be meaningful if P^* and Q^* lie in a vector space. Accordingly we can speak of a linear mapping from any vector space to any vector space. This last remark, made several chapters earlier, would have suggested merely examples of the kind considered at the beginning of Chapter Three, mappings from two dimensions into three dimensions and suchlike. But Chapter Seven, with its variety of vector spaces, shows that linear mappings have a much wider scope; they include differentiation, integration, the operators Δ and E of Chapter Seven, operations in mathematical physics, and many other things besides.

For example, functions $f(x)$ continuous for $0 \leqslant x \leqslant 1$ form a vector space. For every such function, the area under the graph $y = f(x)$ is given by $\int_0^1 f(x)dx$. You tell me the function, $f(x)$; I will tell you A, the area under it. So here we have a mapping, $f(x) \rightarrow A$; input, continuous function; output, the number A. Basic formulas of calculus establish that this mapping is linear (compare the footnote on page 112, establishing that differentiation is a linear mapping). So this definite integral establishes a linear mapping from the space of continuous functions to the space of real numbers.

Instead of considering the area under the graph from 0 to 1 we might consider the area from 0 to x. The result of course depends on x; our output is now not a single number but a function of x. For example, if we began with the graph $y = x^2$ our formula for the

area from 0 to x would be $A(x) = \frac{1}{3}x^3$. The mapping involved here is also a linear one: this is assured by theorems such as 'the integral of a sum is the sum of the integrals'. But the mapping is now $f(x) \rightarrow A(x)$; continuous function $\rightarrow$ continuous function. It is a mapping of the vector space of continuous functions to itself. Accordingly the operation can be repeated; we can consider its square, cube, and other powers; we can combine these to form polynomials in the manner of Chapter Three.

It is, of course, in no way necessary that the continuous functions involved should correspond to simple formulas, as in the example used above, $x^2 \rightarrow \frac{1}{3}x^3$. The input function could be quite irregular, such as, say, a record of barometric pressure.

Not every continuous function can be differentiated (though it took about two centuries to discover this). Accordingly the differentiation operator, $D = d/dx$, cannot be applied to every continuous function $f(x)$. So D does not define a mapping for the vector space of continuous functions. Many of the functions met in the theory of differential equations, such as e^x, $\sin x$, $\cos x$, xe^{-x}, belong to the class of entire functions, specified at the end of Chapter Seven. An entire function can always be differentiated, and the result is an entire function, so the mapping $D; f(x) \rightarrow f'(x)$, maps the space of entire functions to itself. As already observed in the footnote on page 112, D is a linear transformation.

Since D is linear, any polynomial in D will also be linear. If $v = u'' + u$, which means $v = (D^2+1)u$, the mapping $u \rightarrow v$ is a linear one.

AN ELECTRICAL MAPPING

Figure 68 shows a very simple example of the kind of electrical problem considered on page 131. The outside points, A to H, are connected to batteries, and have known potentials $a, b, c \ldots h$. The potentials at the interior points W, X, Y, Z are w, x, y, z; these are not given, but can be determined from the requirement that the potential at each point is to be the average of the potentials at the four neighbouring points (see page 132). This requirement gives us four equations for four unknowns. By solving the equations we find

$$y = (2a+7b+7c+2d+2e+f+g+2h)/24,$$

and very similar expressions for x, w, z. It will be noticed that the expressions for w, x, y, z are linear in $a, b, c \ldots h$; *we therefore have a linear mapping from the potentials chosen on the boundary to the resulting potentials in the interior.*

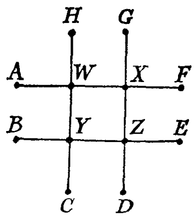


Figure 68

If we considered a finer network with more points involved, our equations would be longer and there would be more of them, but the mapping would still be linear. This remains true even if we go to the limit and consider a piece of continuous copper sheet. The sentence in italics above still holds good.

AN INCONSPICUOUS MAPPING

A linear mapping we very easily overlook is that which is used whenever we substitute in a formula. This is a mapping, function $\rightarrow$ number. You tell me a function; I answer the value of that function when a fixed number, say 10, is substituted in it. The table below shows the effect of substituting 10 in various expressions:

Expression	Value for $x = 10$
x^2	100
x^3	1,000
x^3+x^2	1,100
$5x^2$	500

The mapping, from the first column to the second, is linear. $x^2 \rightarrow 100$, $x^3 \rightarrow 1{,}000$; the sum x^3+x^2 yields, as it should, the sum $1{,}000+100$. In the same way $x^2 \rightarrow 100$; five times x^2 yields five times 100. Thus both tests for linearity are passed.

A linear mapping is involved every time we make tables. Suppose we make tables to show the values of various functions for $x = 2$, $x = 3$, $x = 4$. For example, the table for x^2 would show the three entries 4, 9, 16. We write $x^2 \rightarrow (4, 9, 16)$. In the same way $x^3 \rightarrow (8, 27, 64)$. Here we have a mapping in which each expression yields a vector in three dimensions. It is easily checked that the mapping is a linear one.

Are we just being pompous about a rather simple matter? Does it do us any good to have observed that the mapping is linear? The great advantage of problems involving linear mappings is that it is very easy to fit together solutions of simple problems to obtain the solution of a more complicated one.

Consider the problem: what quadratic expression px^2+qx+r has the value a when $x = 2$, b when $x = 3$ and c when $x = 5$?

What would be a particularly simple problem of this type? If we take $a = 1$, $b = 0$, $c = 0$ the question becomes: what quadratic has the values 1, 0, 0 corresponding to 2, 3, 5 respectively? This question is easier to answer than the general one. For a quadratic that gives nought for $x = 3$ and $x = 5$ must be of the form $k(x-3)(x-5)$. To get the value 1 for $x = 2$ we must take $k = \frac{1}{3}$.

By considering in the same way the special cases $a = 0$, $b = 1$, $c = 0$ and $a = 0$, $b = 0$, $c = 1$ we are led to the three following results:

$$\begin{aligned}(x-3)(x-5)/3 &\rightarrow (1, 0, 0)\\ -(x-2)(x-5)/2 &\rightarrow (0, 1, 0)\\ (x-2)(x-3)/6 &\rightarrow (0, 0, 1).\end{aligned}$$

We now return to our original question; we are looking for a quadratic that yields (a, b, c). Now (a, b, c) can be expressed as a combination of the outputs in the three special cases just considered, for:

$$(a, b, c) = a(1, 0, 0)+b(0, 1, 0)+c(0, 0, 1).$$

But with a linear process, if you want to multiply the output by a, you multiply the input by a; if you want to add outputs, you add inputs. Accordingly, if we take a times our first quadratic, b times our second quadratic, c times our third quadratic, and add these together, we shall obtain the expression we are looking for.

This type of problem was considered at greater length in Chapter Four of *Prelude to Mathematics*, but without any explicit reference to linearity.

Linearity is involved in countless branches of physics – for instance, in the idea that the potential for two electric charges can be found by adding the potentials for the separate charges. Linearity is usually the basis for anything described as a Principle of Superposition.

NOTE ON 'FUNCTIONS'

What we have called a *mapping* is often called a *function.* The word 'function' has a long history, in the course of which its meaning has changed at least twice. To begin with $y = f(x)$ meant that y was

Figure 69

related to x by some simple formula. A graph such as that in Figure 69, consisting partly of a line and partly of a parabola, was not accepted as representing a function. Rather it was considered to show pieces of two different functions. Various infinite series, however, were accepted as defining functions. After furious controversies in the period 1730–60, it gradually became recognized that a single series, a Fourier series involving sines and cosines, could represent the graph in Figure 69, and indeed Fourier series could produce far more violent mixtures of shapes than had ever been envisaged before. The distinction between curves given by a single formula and those given by several could no longer be maintained. It became accepted that $y = f(x)$ could be written if there was any rule, however complicated, that fixed y once x was known.

For example the rule might be that y was $+1$ if x was an even whole number, y was -1 if x was an odd whole number, and y was 0 for all other values of x.

A certain restriction survived; the rule had to fix the *number y* once the *number x* was known. But there are problems, particularly in the calculus of variations, where the unknown is not a number but a curve or a surface. On the surface of an egg, which curve gives the shortest route from A to B? If a soap bubble is bounded by a wire rim, the bubble settles down in the surface that has the least possible area; what surface will that be? These problems have a certain analogy with the maximum–minimum problems in beginning calculus; only instead of having to find a number that makes something a minimum, we have to find a curve or a surface that does so.

Now, given a smooth curve, its length is fixed; given a smooth surface, its area is fixed. This suggests that the essential feature of a function is involved here, and that we may start writing $s = f(C)$ where s is the length of the *curve* C, or $A = f(S)$ where A is the area of the *surface* S. Functions of this kind were considered in 1889 by Arzela and by Volterra. Indeed Arzela's paper in the *Rendiconti Lincei* had the title 'Funzioni di linee', which is best translated 'functions of curves'. In 1903, Hadamard published a paper on 'functional operations'; these were mappings, function $\rightarrow$ number. Definite integration, as used in the first example in this chapter, would be a very simple example of a functional.

The stage had now been reached where the input was no longer a number, but the output still was. It was easy to take the final step, and ask, 'With $y = f(x)$ or $x \rightarrow y$, why do we not let x and y stand for any objects whatever?' Thus, in our example with soap bubbles, provided only that the shape of the wire rim uniquely determines the shape of the soap bubble that spans it, we may write $S = f(R)$ where R denotes the shape of the rim, and S the shape of the resulting soap bubble.

It is conceivable that in 1700 someone might have had the idea of studying functions involving any kind of objects whatever. Even so, he might not have made much progress. To make important discoveries, you do not need only a good question. You need plenty of examples to suggest to you the kind of result that waits to be

discovered. By 1900, enough was known about the calculus of variations, integral equations, and other branches of analysis to make a general theory of functions fruitful.

In classical analysis, one of the first things we ask about a function is whether it is continuous; if $y = f(x)$, does a small change in x lead to a small change in y? In the general theory, with x and y representing arbitrary objects, we naturally try to define continuity. For example, is the function rim → bubble continuous? We have to clarify the question. What does it mean to ask whether a small change in the rim produces a small change in the shape of the bubble?

We might go even further and ask whether we can generalize the ideas of calculus. If x and y are objects other than numbers, can we obtain an equation $dy = f'(x)\,dx$ and give a sensible meaning to it? M. Fréchet in 1925 showed that in certain circumstances this could be done.*

THE FRÉCHET DERIVATIVE

Fréchet's work was extremely clever, and like much clever work it depended on the ability to recognize and to isolate a fairly simple idea. Fréchet looked at what we do when we begin calculus, and

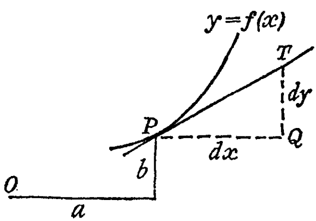


Figure 70

found a way of describing it that remained meaningful and helpful in much more general situations. Figure 70 shows the situation we consider when we first learn to differentiate. The curve $y = f(x)$ passes through the point P with coordinates (a, b); PT is the

* 'La Notion de differentielle dans l'analyse générale', *Annales scientifiques de l'École Normale Supérieure*, vol. XLII, 1925, pp. 293–323.

tangent at P. In this figure we can see that, near the point P, the curve $y = f(x)$ and the tangent PT are almost indistinguishable. The tangent is thus a line that gives an excellent approximation to the curve near P. The theory of differentiation sets out to answer two questions, (1) does such a straight-line approximation to the curve exist? (2) if so, what is it? In elementary calculus, the emphasis is on the question (2); we try to compute the slope of the tangent. Later on, in the more sophisticated approach of analysis, our attention is drawn to the importance of (1). We meet curves that are so crinkly that no line can give even an approximation to their behaviour.

In Figure 70 the lengths of PQ and PT have been shown as dx and dy. Thus (dx, dy) are the coordinates of a point T on the tangent with respect to an origin placed at P. There is no suggestion here that dx and dy are 'infinitesimally small'. For example, if the graph $y = f(x)$ happened to be $y = x^2$ and P the point (1, 1), we should have $dy = 2dx$. It would be perfectly in order to put $dx = 3$, $dy = 6$ and conclude that the point 3 to the east and 6 to the north of P lies *on the tangent*. The idea of smallness arises only *if we want to use the tangent as an approximation to the curve*. The tangent is close to the curve only when we are near to P. If we put $dx = 0{\cdot}001$ and $dy = 0{\cdot}002$, we can conclude that the point whose coordinates (in the original system, with origin at O) are $(1{\cdot}001, 1{\cdot}002)$ lies *exactly* on the tangent and looks like being *approximately* on the curve.

Accordingly the tangent at P may be defined as the line that gives the best approximation to the graph $y = f(x)$ near P. The business of differentiation is to find this best *linear approximation*, when it exists.

Now in this chapter we have considerably enlarged the meaning and scope of the word *linear*. The formulation just given enables us similarly to enlarge our idea of differentiation. In any situation where we have a complicated mapping, we can ask whether there is a linear mapping that gives a good approximation to it over a limited region, and, if so, what it is. A theory that answers these questions will be called a theory of differentiation. The analogy with ordinary calculus will help to suggest both results in this theory and ways of applying these results.

In ordinary calculus, the linear approximation can often be calculated by simple algebra. Consider the example already used, the behaviour of the curve $y = x^2$ near the point (1, 1). If we put $x = 1+h$, we find $y = 1+2h+h^2$. When h is small, h^2 is much smaller. If h is a few thousandths, h^2 is a few millionths. If we agree to neglect h^2, we find the point $x = 1+h$, $y = 1+2h$ to be close to the curve. Giving different small values to h, we obtain points that lie in a line; we have thus found a linear approximation to the curve, the tangent at P. We may write $dx = h$, $dy = 2h$ and deduce $dy = 2dx$, the linear equation that specifies the tangent.

The procedure, crudely stated, is to neglect all except the linear terms; this, not surprisingly, leads to a linear relation. The same procedure can be followed in less familiar situations. Suppose, for instance, we are dealing with a mapping from two dimensions to two dimensions, $(x, y) \rightarrow (u, v)$, where $u = 3x^2+y^2$ and $v = xy$. We are interested in what happens, say, near $x = 1$, $y = 2$. If we put $x = 1+h$, $y = 2+k$, we find

$$\begin{aligned} u &= 7+6h+4k+ \dots \\ v &= 2+2h+k+ \dots \end{aligned}$$

where the dots indicate terms that would be measured in millionths if h and k were measured in thousandths. These terms are to be neglected when we make our approximation. We thus arrive at the linear approximation

$$\left.\begin{aligned} du &= 6\,dx+4\,dy \\ dv &= 2\,dx+\,dy \end{aligned}\right\}. \tag{1}$$

If we write dU for (du, dv) and dX for (dx, dy) we can express the equations (1) above in the compact form $dU = M\,dX$, where M denotes the matrix $\begin{pmatrix} 6 & 4 \\ 2 & 1 \end{pmatrix}$. In this situation, the matrix M plays the role of a differential coefficient.

MAXIMUM AND MINIMUM

One of the problems considered in calculus is to find where a maximum or minimum occurs. The argument runs, very roughly,

like this. Suppose that on the graph $y = f(x)$ we have a point P where the tangent PT is uphill, as in Figure 71. Then P cannot be a maximum. For the tangent is rising, so by going to the right, we can find a point M, on the tangent, that is higher than P. Now the tangent gives a good approximation to the curve, so if M is sufficiently close to P, there will be a point R, on the curve, very close indeed to M and accordingly also higher than P. Thus P cannot be the highest point of the curve in this region. Nor can it be the lowest. By going to the left, we find a point, L, on the tangent and

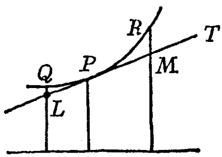


Figure 71

lower than P. Very close to L is Q, on the curve and lower than P. So P cannot be the lowest point in the region.

This argument clearly needs some tidying up. In the figure, the point Q is higher than L. In a detailed proof, we should have to show carefully that Q cannot be so far above L as to be actually higher than P. This point can be satisfactorily cleared up, but we do not wish to enter into it now.

A very similar argument shows that a maximum or minimum cannot occur at a point where the tangent is downhill.

So the only place where a maximum or minimum can possibly occur is where the tangent is horizontal.* This is the basis of the usual procedure, where we search for maxima and minima by solving $f'(x) = 0$ and then applying certain further tests.

We are going to adapt this procedure first of all to a problem in three dimensions, and then try just to outline considerations show-

*We are ruling out such things as pointed peaks, where $f'(x)$ does not exist, and also problems in which the domain of definition is restricted, e.g. find the maximum of x^2 given that x is positive and does not exceed 10. The maximum occurs for $x = 10$, where the slope of the tangent is not 0 but 20.

ing that it can be applied to a problem involving an infinity of dimensions. One might anticipate that this last project would call for some kind of radical new thinking. This is not so; the difficulty, if any, arises from the fact that a certain amount of traditional calculus symbolism and manipulation is called for. Any reader who is out of touch with such work is advised simply to skim over the remainder of this chapter, and be content to gain a general impression of the kind of problem to which the Fréchet derivative can be applied.

First we consider the situation in three dimensions. As a particular example we may consider the cuplike surface $z = x^2+y^2$. It can be shown that the point $(x+dx, y+dy, z+dz)$ will lie on the tangent plane at (x, y, z) if $dz = 2x.dx+2y.dy$. More generally, we can consider any surface $z = f(x, y)$ with a tangent plane at the point P given by $dz = p.dx+q.dy$. If $p = 0$ and $q = 0$, we shall have $dz = 0$ whatever numbers are chosen for dx and dy. This means that the tangent plane is horizontal. We want to show that this is the only case in which a maximum or minimum can occur; that if the tangent plane is tilted there will be some points on it higher than P and other points lower than P. Geometrically this appears reasonable, and algebraically there is a little trick that proves it immediately. If we take $dx = p$ and $dy = q$ and substitute in $dz = p.dx+q.dy$ we find $dz = p^2+q^2$, which is bound to be positive, and so gives a point higher than P. Similarly, by taking $dx = -p$ and $dy = -q$, we find $dz = -p^2-q^2$, which is certainly negative and means we have found a point lower than P. Thus we have established that there are uphill and downhill directions on the tangent plane. If we go a small distance in one of these directions, since the surface clings closely to the tangent plane, we should be able to find points of the surface itself that are higher than P and points that are lower than P.

Accordingly, it is impossible for a maximum or minimum to occur except where $p = 0$ and $q = 0$.

In our particular example, $z = x^2+y^2$, we have $dz = 2x.dx+2y.dy$ so $p = 2x$ and $q = 2y$. The condition $p = q = 0$ thus means $x = 0, y = 0$. This in fact gives a minimum, the point at the bottom of the cup, as indeed, in this particular example, might very well have been guessed without the help of calculus.

SPACE OF INFINITE DIMENSIONS

Continuous functions, as we saw at the end of Chapter Seven, form a space of infinite dimension. We now consider a mapping of the type: continuous function $\rightarrow$ number. Let $y = f(x)$ be the graph of a continuous function, defined for x from 0 to 1. Suppose a number, a, defined by the equation:

$$a = \int_0^1 (2xy - y^2)\, dx. \tag{2}$$

The mapping $f(x) \rightarrow a$ resembles some of the mappings we considered near the beginning of this chapter; you tell me what function $f(x)$ you have chosen, I will answer the number a that results when $y = f(x)$ is substituted in equation (2). The table below gives a few of the functions you might choose, and the resulting values for a.

$y = f(x)$	a
$y = 0$	0
$y = \frac{1}{2}$	0·25
$y = 1$	0
$y = x$	0·3333 . . .
$y = 1+x$	−0·6666 . . .
$y = x^2$	0·3

Observe first of all that this mapping is *not* linear. If it were, the number corresponding to $y = 1+x$ could be found by adding the numbers corresponding to $y = 1$ and $y = x$. The number corresponding to $y = 1$ would be twice the number corresponding to $y = \frac{1}{2}$. Neither is in fact so.

We might be interested in finding which function produced the largest value of a. In the table above, the largest value, 0·3333 . . ., occurs for $y = x$. Is this in fact the maximum, or would some other choice of y, not shown in the table, produce an even larger number?

Suppose this question had been posed after the first three rows had been calculated. At this stage $y = \frac{1}{2}$ had given the largest number, 0·25. On the analogy of our earlier calculus work, we might see whether some graph that differed very little from $y = \frac{1}{2}$

would give a value just a little larger than 0·25. If we could do this, we would have demonstrated that there was an 'uphill' direction and that we could not be dealing with a maximum. Accordingly, we will consider $y = \frac{1}{2} + \phi(x)$, where it is understood that for each x the value $\phi(x)$ is quite small. The graphs might then be as shown in Figure 72.

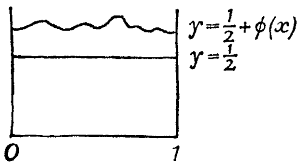


Figure 72

When we substitute $y = \frac{1}{2} + \phi(x)$ a brief calculation leads us to the number

$$0{\cdot}25 + \int_0^1 (2x-1)\phi . dx - \int_0^1 \phi^2 \, dx.$$

We want to make a simplifying approximation here, so we once more use the argument that if ϕ is measured in thousandths, then ϕ^2 is measured in millionths, and may be expected to be relatively unimportant. Accordingly, to obtain a linear approximation, we neglect the last term, $\int_0^1 \phi^2 dx$. The first term, 0·25, is simply the number corresponding to $y = \frac{1}{2}$. Thus the approximate change in the number is given by the second term. Changing y from $\frac{1}{2}$ to $\frac{1}{2} + \phi$ makes *a* change from 0·25 to $0{\cdot}25 + \alpha$ approximately, where:

$$\alpha = \int_0^1 (2x-1) \, \phi . dx. \tag{3}$$

Note that the mapping $\phi \to \alpha$ is linear. The equation (3) is thus similar to the approximations we found earlier – the tangent line as an approximation to a small piece of a curve, the tangent plane as an approximation to a small piece of a surface.

Can we go 'uphill'? Is there any small change ϕ that will make α positive? There certainly is. If we take $\phi(x) = 0{\cdot}001\ (2x-1)$ we find:

$$\alpha = 0{\cdot}001 \int_0^1 (2x-1)^2\, dx$$

which, containing a square as it does, must be positive.

It can in fact be tested by calculation that the approximations we have made are justified and that $y = \frac{1}{2}+0{\cdot}001\ (2x-1)$ does give a slightly larger value for a than $y = \frac{1}{2}$ produces. There is an uphill direction and $y = \frac{1}{2}$ cannot give a maximum.

We now reflect that the considerations used above could be applied not only to $y = \frac{1}{2}$ but to any equation $y = f(x)$ that was thought likely to yield a maximum. We will suppose that replacing $f(x)$ by $f(x)+\phi(x)$ causes the corresponding number to change from a to $a+\alpha$ approximately, the approximation being made by neglecting ϕ^2. By a calculation, along exactly the same lines as that already done, we shall find, instead of equation (3), the equation:

$$\alpha = \int_0^1 2(x-f)\ \phi\ .\ dx. \tag{4}$$

We can now play the same trick as before. Taking $\phi = 0{\cdot}001\ (x-f)$ we find that $\alpha = \int_0^1 0{\cdot}002\ (x-f)^2\ dx$. Here again we have an expression involving a square; it cannot be negative and for a moment we may think we have shown that no maximum exists at all. It looks as if, whatever $f(x)$ we chose, we could always make a small change in it that would increase the corresponding number. However, this argument overlooks one possibility. If $f(x) \equiv x$, then $x-f = 0$ and equation (4) shows that $\alpha = 0$ for every possible ϕ. This is the situation we meet at the flat top of a hill; every tangent is level, there is no way up. What our argument does show is that this situation can arise *only* for $y = x$.

At the beginning of calculus we search for maxima and minima first by determining where the tangent is horizontal, and then by applying further tests to see whether we have found a maximum, a minimum, or merely a point of hesitation (a horizontal inflexion). It should be clear that the work we have just done corre-

sponds only to the first part – we have looked for level tangents. I have implied that $y = x$ corresponds to a hilltop, but you have only my word for this; I have not here discussed how we distinguish between a hilltop, a valley bottom, and a mountain pass.

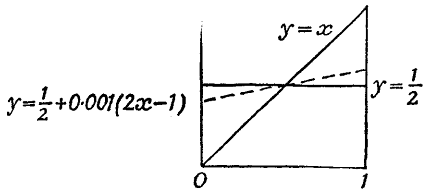


Figure 73

Figure 73 shows the original suggestion $y = \frac{1}{2}$, the improved suggestion $y = \frac{1}{2} + 0{\cdot}001\ (2x - 1)$ and the final solution $y = x$. It may be noted that the amendment made to $y = \frac{1}{2}$ takes us in the direction of the final solution. The graph $y = \frac{1}{2}$, as compared with the solution $y = x$, is too high in the first half of the interval and too low in the second half. The amendment tends to correct these faults; it brings down the earlier part and raises the latter part of the graph. This suggests a method that can be used in some practical problems where no simple theoretical solution is known. The recipe is simple – start anywhere and keep going uphill. There is no guarantee that this will bring you to the top of the highest peak, but with luck it will bring you to the top of some hill. If the procedure is repeated with a variety of different starting points, one can get at any rate some indication of where the mountains lie.

CHAPTER NINE

What is a Rotation?

In Chapter One we explained affine geometry to a disembodied spirit. In this chapter, we try to do the same with Euclid's geometry. We start with the geometry of the plane. The work here is elementary, using only Pythagoras' Theorem and the elements of algebra – in particular, the formula for $(a-b)^2$. This part of the chapter could be used with fairly young pupils, and it is here presented very much as it might be done in school by the discovery approach, the class always being asked, 'What shall we do now?'

Later, we examine this early work and see whether it can be generalized to n dimensions.

This chapter helps to show how a mathematical subject reaches abstract form. Anything we say to the spirit is of necessity abstract; we cannot draw pictures. But we can and do draw pictures while we are discussing among ourselves what message to send. The spirit is outside the subject; we are inside it. This is as it should be. The impression is sometimes given that abstract work means not knowing what you are talking about. This is quite false. A mathematician seeking to create an abstract subject begins with various situations which are perfectly familiar to him; he knows exactly what they are; he then tries to extract their common features and to build a single theory that covers these. The theory is abstract only because it omits any mention of points in which the situations differ.

Now we begin our exposition of plane geometry to the spirit. We draw on our experience of graph paper and begin by defining 'point'.

1. *Definition.* A point P is something specified by a pair of numbers (x, y).

Next we want to define the distance between two points, P, (x_1, y_1), and Q, (x_2, y_2). How shall we do this? A class of children would probably begin with numerical examples, and gradually extract the general argument. In Figure 74, the distance PQ is

If A^* is the point specified by (c, s), we obtain the equation

$$c^2+s^2 = 1. \tag{1}$$

This is all we can say about A^*. Now let us consider what happens to B, $(0, 1)$. All distances are to stay the same, so the movement must not change the distances OB and AB. It is as if B were tied down by two bars, as in Figure 76. Suppose B^* is the

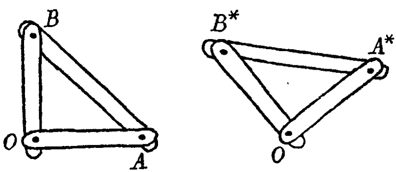


Figure 76

point (p, q). The conditions $O^*B^* = OB$ and $A^*B^* = AB$ lead to equations (2) and (3):

$$p^2+q^2 = 1, \tag{2}$$

$$(p-c)^2+(q-s)^2 = 2. \tag{3}$$

These equations for the unknowns p, q do not look particularly simple. However, it often happens that, in studying a problem that arises naturally, we are helped on our way by a series of unexpected simplifications. This happens here. When we multiply equation (3) out, we find it contains p^2+q^2, which by (2) is simply 1, and c^2+s^2, which by (1) is also 1, and the equation boils down to:

$$cp+sq = 0. \tag{4}$$

If $s \neq 0$,* we may deduce $q = -cp/s$ and substitute in (2). This leads, after clearing of fractions, to $p^2(c^2+s^2) = s^2$. Once again we have a simplification, for $c^2+s^2 = 1$. We find $p^2 = s^2$ so $p = s$ or $-s$. The value of q follows from $q = -cp/s$. So we have

* If $s = 0$, $c \neq 0$ and we obtain exactly the same solution by using $p = -qs/c$. By a more sophisticated handling of the algebra, one can avoid this separation of alternative possibilities.

would apply to any translation. So we reach our next two messages:

4. *Definition.* Any mapping of the form $x^* = x+a$, $y^* = y+b$ is called a *translation.*

5. *Theorem.* Every translation is a rigid movement.

MOVEMENTS ABOUT THE ORIGIN

Are there other rigid movements besides translations? We naturally expect there will be – namely, rotations. Imagine a piece of cardboard sliding on a flat table. We can make translations impossible by sticking a pin through the card into the table. The card would still be free to turn. We may suppose the pin stuck in at the origin. This leads us to our next definition.

6. *Definition.* A *movement about the origin* is a mapping such that $O^* = O$.

We now start looking for rigid movements that leave the origin fixed, $O^* = O$.

Consider the point A, (1, 0). It goes to A^*. Can A^* be chosen anywhere we like? No, for this is to be a rigid movement. That means $O^*A^* = OA$. We can imagine O and A connected by a steel bar, which must not be stretched, compressed or broken. We picture the situation as in Figure 75; A^* must be somewhere on the circle, centre O, radius 1.

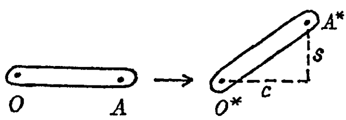


Figure 75

For the spirit, we merely need quote that a rigid movement does not change distances, so we must have $O^*A^* = OA$. As both of these involve square roots (Definition Two), it will be more convenient to write $(O^*A^*)^2 = (OA)^2$; in the later work, this squaring will be convenient every time two lengths are equated; we shall not mention it on each occasion.

brings us naturally towards the idea of non-Euclidean geometry and of metric spaces in general. (A metric space is one where distance has been defined in some reasonable way.)

In fact, we should soon notice it if the formula for distance were replaced by one chosen at random. For instance, in a universe where, instead of $PQ^2 = (x_2-x_1)^2+(y_2-y_1)^2$, we had $PQ^4 = (x_2-x_1)^4+(y_2-y_1)^4$, we should find we were unable to turn round. Whatever direction we were born facing, we should continue facing for the rest of our lives. The reason will appear in the course of this chapter (see page 175).

Our next message to the spirit will in fact be concerned with how rigid objects move. In class discussion, we would demonstrate things that can be done to a flexible object, such as a handkerchief or a rubber tube, but cannot be done to a rigid body, such as a brick or a steel bar. In what does the difference lie? From discussion it should emerge that a rigid body can only move in a way that keeps the distance between any two points of it unchanged.

In any movement, a point goes from its starting position P to its final position P^*. This is a mapping, $P \to P^*$. So we arrive at our next definition.

3. *Definition*. A movement is said to be rigid (in technical jargon, an isometry) if it preserves distances; that is, if $P \to P^*$ and $Q \to Q^*$, then $PQ = P^*Q^*$ always.

There is always one rigid mapping – the identity mapping; leave every point where it was. It is not obvious (to the spirit) that any others will exist. There are geometries in which no genuine movement is possible. The surface of a statue is such a geometry. If a statue is covered by a tightly fitting suit of chain armour, we would not expect to be able to slide the armour about on the statue. It would fit in one position only. For all the spirit knows, our Definition Two may describe such a situation. It is no use our saying, 'Obviously you can slide any object resting in contact with a plane.' We have to prove that rigid movements exist.

We met a rigid movement in Chapter One. Figure 4 showed the effect of the translation $x^* = x+2$, $y^* = y+1$. We can prove that this is a rigid movement. For $x_2{}^*-x_1{}^* = x_2-x_1$ and $y_2{}^*-y_1{}^* = y_2-y_1$. Using the formula in Definition Two to write down PQ and P^*Q^*, we see at once that $PQ = P^*Q^*$. The same argument

given by Pythagoras' Theorem, $PQ^2 = PM^2 + MQ^2$. This leads to $PQ^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2$. So our next message to the spirit runs:

2. *Definition.* The distance, PQ, from the point P, (x_1, y_1) to the point Q, (x_2, y_2) is:

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

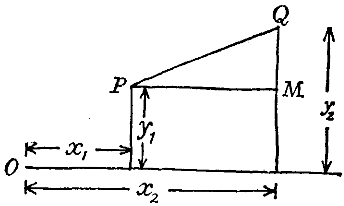


Figure 74

The spirit's reaction to this would probably be, 'What an extraordinary universe you live in! Why, of all the possible formulas, is this strange one chosen for distance?'

Now, of course, if we were just starting to study geometry, we should never have thought of this definition. Two thousand years of geometrical experience lie behind it. This is an important thing to realize about the formal, abstract, axiomatic, deductive aspects of mathematics – these represent a final stage of the subject, rarely a beginning. Note too how the order of development changes. Pythagoras' Theorem was the forty-seventh proposition of the first book of Euclid; it was the goal and climax of that book. Here it is Definition Two.

As a rule, we take distance for granted and do not feel any need to define it. The reason for the abstract approach (that is, trying to express things in a way intelligible to a disembodied creature) is to force ourselves to think about what we usually take for granted. And it leads us immediately to the very significant question asked by the spirit, 'What would the universe be like if it had been built on some formula other than the one in Definition Two?' This

two solutions: Solution 1, $p = -s, q = +c$; Solution 2, $p = +s$, $q = -c$.

If we examine these solutions graphically, as in Figure 77, we

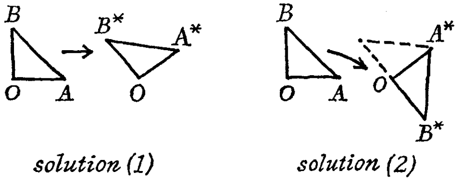


Figure 77

find that in (1) the triangle OAB has simply been rotated, in (2) the triangle has been rotated and then turned over.

It may well be objected that 'rigid movement' is not an accurate description of (2), since we could not move the triangle from position OAB to $O^*A^*B^*$ without taking it outside the plane of the paper. One could avoid this by talking about 'isometries' (mappings that preserve distances) instead of 'movements'. This removes the objection, but at the cost of replacing a simple, familiar word by a piece of unfamiliar jargon. Each person must weigh the advantages and disadvantages for himself.

Whatever name we decide to use, we want to keep Solution 2, for Euclid's theorems are not damaged by reflection in a mirror.

We will develop Solution 1 in detail. A similar development of Solution 2 is possible, and may be done as an exercise.

Accordingly, we are now supposing that A goes to A^*, (c, s), and that B goes to B^*, $(-s, c)$, the position given by Solution 1.

Let us now consider what happens to any point P with coordinates (x, y). As Figure 78 shows, P can be thought of as joined to O, A, and B by bars. None of these may change in length and in consequence the two coordinates (x^*, y^*) of P^* must satisfy the three conditions $OP^* = OP$, $A^*P^* = AP$, $B^*P^* = BP$. Now in general, three equations cannot be satisfied by two unknowns. This is the reason why we should not take rotations for granted. In most geometries, no such things exist. Definition Two uses one of

the very few formulas that permit continuous rotation of a rigid object.

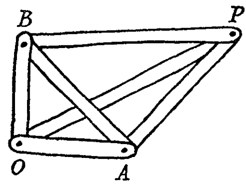


Figure 78

The three conditions lead to equations (5), (6), (7):

$$(x^*)^2+(y^*)^2 = x^2+y^2, \tag{5}$$

$$(x^*-c)^2+(y^*-s)^2=(x-1)^2+y^2, \tag{6}$$

$$(x^*+s)^2+(y^*-c)^2 = x^2+(y-1)^2. \tag{7}$$

Here x, y, c, s are supposed given; x^* and y^* have to be found. Once again, a considerable simplification occurs. If equations (6) and (7) are multiplied out, and use is made of (1) and (5), we are led to equations (8) and (9):

$$x^*c+y^*s = x, \tag{8}$$

$$-x^*s+y^*c = y. \tag{9}$$

If we solve these equations for x and y, once again using the fact that $c^2+s^2 = 1$, we find:

$$\left.\begin{aligned} x^* &= cx-sy \\ y^* &= sx+cy \end{aligned}\right\}. \tag{10}$$

It is now necessary to check that these equations do in fact give a solution to all three equations (5), (6), and (7). We are pleased (and the spirit is surprised) to find they do.

But we are still not finished. Checking equations (5), (6), and (7) shows that the distances of any point from O, A, B remain unaltered. But the spirit might make the following objection. Suppose we apply the transformation (10) to two points, P with co-

ordinates (x_1, y_1) and Q with coordinates (x_2, y_2). The work we have done guarantees that the distances OP, AP, BP will be preserved unchanged, and similarly for OQ, AQ, BQ, but nothing has been said about the distance PQ, indicated by the dotted line in Figure 79. On the basis of our geometrical experience, we are

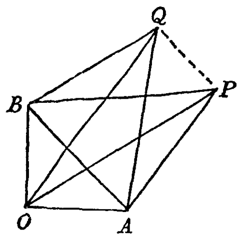


Figure 79

pretty confident that P^*Q^* will in fact turn out equal to PQ, but this is no help to the spirit. To convince the spirit, we will have to calculate P^*Q^* and verify that it equals PQ – which in fact it does.

All this work has been on the basis of Solution 1. A similar development of Solution 2 would also lead to an isometry, with the equations:

$$\left.\begin{aligned} x^* &= cx + sy \\ y^* &= sx - cy \end{aligned}\right\}. \tag{11}$$

It is noticeable that both (10) and (11) are *linear transformations*. On the basis of all this algebra, we can announce the following results.

7. *Theorem.* There are an infinity of rigid movements (isometries) that leave the origin fixed. They are of two types, type 1 being specified by equations (10) and type 2 by equations (11).

8. *Theorem.* Every rigid movement about the origin is given by a linear transformation.

ROTATIONS AND REFLECTIONS

The great difference between transformations of type 1 and type 2 is clear to us from Figure 77. But the spirit cannot look at this figure and does not know what we mean when we talk about 'turning over'. Is there any way in which we could convey something of this difference to the spirit?

In Chapters Four and Five, we saw that a good question to ask about a linear transformation was, 'What are its eigenvectors?' This question was answered with the help of the characteristic equation.

Now as a rule a rotation changes the direction of every vector; we do not expect to find any eigenvectors for a rotation. If R denotes $\begin{pmatrix} c & -s \\ s & c \end{pmatrix}$, the matrix corresponding to equations (10), we find the characteristic equation of R is $R^2-2cR+(c^2+s^2)\,I = O$. The roots corresponding to this equation are $c \pm is$, where $i = \sqrt{-1}$. Except when $s = 0$ (rotation of $0°$ or $180°$) these roots are complex numbers, and it can be shown (as we guessed) that there is no real vector unaltered in direction by the rotation R. Incidentally, if we allow complex numbers, we find that all rotations have the same eigenvectors, namely, $(1, i)$ and $(1, -i)$, a famous and surprising result.*

Things are far otherwise when we come to transformations of type (2). Writing M for $\begin{pmatrix} c & s \\ s & -c \end{pmatrix}$, the matrix given by equations (11), we find its characteristic equation to be $M^2 = I$. This equation was studied at some length in the earlier part of Chapter Five. We found that, for any vector v whatever, M leaves $\frac{1}{2}(I+M)v$ unchanged but multiples $\frac{1}{2}(I-M)v$ by -1. If we choose $v = (2k, 0)$, we see that M leaves every vector of the form $k(1+c, s)$ unchanged but multiplies every vector of the form $k(1-c, -s)$ by -1. The vectors – perhaps it would be better to say, the points – of the form $k(1+c, s)$ fill a certain line through the origin; M in fact represents a reflection in this line.

*Compare the final section of *Prelude to Mathematics*, Chapter Eleven, 'The Circular Points at Infinity'.

The symbols R and M were chosen as being the initial letters of *rotation* and *mirror*.

The distinction we have drawn between R and M is one that can be appreciated by the spirit. A rotation R leaves the one point O where it was. A reflection M leaves each point of a line where it was.

The spirit could also appreciate the different ways in which rotations and reflections combine. A rotation followed by a rotation gives a rotation; rotations in fact form a group. Reflections do not; two reflections produce, not another reflection, but a rotation.

We of course see these results geometrically, but the spirit can test them by matrix multiplication. If $R_1 = \begin{pmatrix} c & -s \\ s & c \end{pmatrix}$ and $R_2 = \begin{pmatrix} C & -S \\ S & C \end{pmatrix}$ we find $R_1R_2 = \begin{pmatrix} cC-sS & -cS-sC \\ sC+cS & cC-sS \end{pmatrix}$. This is of the form $\begin{pmatrix} p & -q \\ q & p \end{pmatrix}$, which is correct for a rotation, with $p = cC-sS$, and $q = sC+cS$. For R_1R_2 to be a rotation, we must also have $p^2+q^2 = 1$. It can be verified that $p^2+q^2 = (c^2+s^2)\ (C^2+S^2)$ and this will be 1, since $c^2+s^2 = 1$ and $C^2+S^2 = 1$ are necessary to make R_1 and R_2 rotations.

It will probably be apparent now why the letters c, s were chosen for the coordinates of A^*. They are the initial letters of *cosine* and *sine*. In our treatment, c and s appeared as the coordinates of the point to which the rotation R_1 sends the point (1, 0). This is in fact a very convenient way of introducing sine and cosine. It holds whatever the angle involved in the rotation may be. We thus avoid the awkwardness of defining sine and cosine first of all for angles between 0° and 90°, and then having to make excuses when we want to extend the definition to angles of any size. It will be noticed that the equations $p = cC-sS$ and $q = sC+cS$ in the previous paragraph express the addition formulas for sine and cosine. It is interesting that we have been able to get so far with trigonometry without (until now) even mentioning the word.

Exercises

1. Given the reflections $M_1 = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$ and $M_2 = \begin{pmatrix} A & B \\ B & -A \end{pmatrix}$ where $a^2+b^2 = 1$ and $A^2+B^2 = 1$, verify that M_1M_2 is a rotation.

2. If $X = M_1M_2$ and $Y = M_2M_1$ are X and Y equal? If not, are X and Y related in any simple way?
3. If R_1 is a rotation, as specified in the text above, what kind of transformation is R_1M_1?

GENERALIZATION

In two dimensions we succeeded in finding all possible rigid movements about the origin by a direct attack using elementary algebra. If we go to three, four, five or more dimensions the work becomes increasingly laborious. It seems wise to analyse our experience with two dimensions and see if we can extract principles that will guide us in the general problem for n dimensions.

Now in two dimensions things worked out more simply than we might have expected. The conditions that had to be met were all expressed by equations involving squares – equations (1), (2), (3), (5), (6), and (7). The transformations eventually found, in equations (10) and (11), however, were linear. How did this come about?

The first equation not involving squares was (4). In equation (3), the left-hand side represented $(A^* B^*)^2$ and was $(p-c)^2+(q-s)^2$. When this was multiplied out, it was found to contain p^2+q^2, which we had already met as $(OB^*)^2$, and $c^2 + s^2$, already known as $(OA^*)^2$. These parts were subtracted, and then a factor -2 removed to give equation (4). This is equivalent to saying that you get equation (4) if you add equations (1) and (2), subtract (3) and then divide by 2. Now equations (1), (2), and (3) stated $(OA^*)^2 = OA^2$, $(OB^*)^2 = OB^2$ and $(A^* B^*)^2 = AB^2$. Equation (4) thus states:

$$\tfrac{1}{2}\left\{(OA^*)^2+(OB^*)^2-(A^* B^*)^2\right\} = \tfrac{1}{2}\left\{OA^2+OB^2-AB^2\right\}.$$

In other words, equation (4) states that the movement does not change the value of $\frac{1}{2}(OA^2+OB^2-AB^2)$. This statement, geometrically more complicated than equations (1), (2) and (3), turns out to be algebraically simpler.

Let us see if the same trick will work in three dimensions. Suppose we have any two points, P, (x_1, y_1, z_1) and Q, (x_2, y_2, z_2). A rigid movement does not alter any distance, so it certainly

will not alter the value of $\frac{1}{2}(OP^2+OQ^2-PQ^2)$. How does this expression look in terms of algebra?

$$OP^2 = x_1^2+y_1^2+z_1^2 \qquad OQ^2 = x_2^2+y_2^2+z_2^2$$
$$PQ^2 = (x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2$$

It is seen that:

$$\tfrac{1}{2}(OP^2+OQ^2-PQ^2) = x_1\,x_2+y_1\,y_2+z_1\,z_2.$$

Accordingly, no rigid movement about the origin can change the value of $x_1x_2+y_1y_2+z_1z_2$ for any pair of points.

The argument is reversible; if a movement does not change the values of OP, OQ, and $x_1x_2+y_1y_2+z_1z_2$, it follows from the last equation we had that it does not change the value of PQ.

Here we seem to have two kinds of condition, one algebraic, the other involving the lengths OP and OQ. But $OP^2 = x_1^2+y_1^2+z_1^2$, which may be written $x_1x_1+y_1y_1+z_1z_1$. This is of the same form as the algebraic expression; it is what $x_1x_2+y_1y_2+z_1z_2$ would become if (x_2, y_2, z_2) were allowed to coincide with (x_1, y_1, z_1). A similar remark applies to OQ^2.

Accordingly, the condition for a rigid movement can be stated; *the value of* $x_1x_2+y_1y_2+z_1z_2$ *must remain unchanged for all points* (x_1, y_1, z_1) *and* (x_2, y_2, z_2) *whether distinct or not.*

SCALAR PRODUCT

This expression $x_1x_2+y_1y_2+z_1z_2$ is so important that it receives a special name; it is called the scalar product of the vectors P and Q. Sometimes it is written $P\,.\,Q$ and called 'the dot product'. We shall write it (P, Q).

The scalar product makes various appearances in mechanics – for example as the work done by the force OP in the displacement OQ.

There is a particular situation in which its geometrical meaning is evident. We saw above that the scalar product (P, Q) arose from the expression $\frac{1}{2}(OP^2+OQ^2-PQ^2)$. Pythagoras' Theorem tells us that, if POQ is a right-angled triangle, this expression will be zero, and so $(P, Q) = 0$. We shall use this result in reverse; we shall *define* 'OP perpendicular to OQ' as meaning $(P, Q) = 0$.

Why is it called a 'product'? The reason is that it has many

properties reminiscent of multiplication. For example, we have $(P, Q) = (Q, P)$; $(P, Q+R) = (P, Q)+(P, R)$; $(kP, Q) = k(P, Q)$. All of these can be verified immediately by going back to the algebraic definition of scalar product. These properties justify us in carrying over to work with scalar products many of our habits from elementary algebra. For example, we can multiply out $(A+B,\ C+D)$ as $(A,\ C)+(A,\ D)+(B,\ C)+(B,\ D)$ or get the 'difference of squares' formula $(A+B,\ A-B) = (A,A)-(B,B)$.

Note carefully that the system here does not pass test (6) of Chapter Seven. We have defined a product $(P,\ Q)$ but it is *a single number, not a vector* $(x,\ y,\ z)$. We are not dealing with a ring. In a ring, an expression such as $ab+c$ is meaningful. But $(P,\ Q)+R$ is nonsense; it shows a number added to a vector.

CHANGING AXES

The scalar product is particularly useful when we are changing from one set of *perpendicular* axes to another. For example, suppose three perpendicular vectors P, Q, R are given; we want to use them as axes, that is, to express some vector V in the form $aP+bQ+cR$. We suppose P, Q, R, and V are all expressed in their $(x,\ y,\ z)$ form, so that we can calculate any scalar products we may require. To find a, all we need do is form the scalar product $(P,\ V)$. For if $V = aP+bQ+cR$, then $(P,\ V) = (P,\ aP+bQ+cR)$ $= a(P,\ P)+b(P,\ Q)+c(P,\ R)$. Now fortunately $(P,\ Q) = 0$ since P is perpendicular to Q; similarly $(P,\ R) = 0$, as P is perpendicular to R. All that remains is $(P,\ V) = a\,(P,\ P)$, and we solve this for a.

An example will show the simplicity of this method. For example, the three vectors $P = (1,\ 1,\ 1)$, $Q = (1,\ 2,\ -3)$, $R = (-5,\ 4,\ 1)$ are perpendicular, as can be verified by finding their scalar products. Any vector in three dimensions can be expressed with these as axes. Suppose we want to do this for $V = (10,\ 4,\ -8)$. Then $(P,\ V) = 6$ and $(P,\ P) = 3$, so the equation $(P,\ V) = a\,(P,\ P)$, proved above, becomes simply $6 = 3a$. Accordingly, $a = 2$. In the same way, one can establish $(Q,\ V) = b\,(Q,\ Q)$; this gives $42 = 14b$, so $b = 3$. Finally we can prove $(R,\ V) = c(R,\ R)$ and this gives $-42 = 42c$, so $c = -1$. Thus $V = 2P+3Q-R$. The correctness of this result can be checked easily.

Exercises

1. Express $V = (7, 7, 7)$ in terms of $P = (-2, 3, 6)$, $Q = (6, -2, 3)$, $R = (3, 6, -2)$.
2. Express $V = (6, 1, 4)$ in terms of $P = (1, 0, -1)$, $Q = (1, 2, 1)$, $R = (1, -1, 1)$.

MOVEMENTS IN THREE DIMENSIONS

If we look back at what we did when we were finding the rigid movements in two dimensions, we shall see that the argument fell into two parts. In the first part we were only concerned about what happened to the points A and B on the axes. Now OA and OB were of unit length and OA was perpendicular to OB. Naturally a rigid movement did not alter these facts; OA^* and OB^* had to be of unit length, and OA^* perpendicular to OB^*. These conditions were expressed by equations (1), (2), (4), and solutions (1) and (2) gave all ways of meeting them.

The rest of the investigation showed that no further restrictions were needed. Provided we had found suitable places, A^* and B^*, for A and B to go to, this determined where every other point had to go, and all the conditions of the problem were met. In fact the resulting transformation had to be linear; if $A \to A^*$ and $B \to B^*$, then $xA+yB \to xA^*+yB^*$ gave the required rigid movement.

In three dimensions, it would seem we ought to consider first what happens to the points A, B, C where $A = (1, 0, 0)$, $B = (0, 1, 0)$, $C = (0, 0, 1)$. The vectors OA, OB, OC are of unit length and mutually perpendicular. Algebraically, this is expressed by $(A, A) = (B, B) = (C, C) = 1$ and $(A, B) = (A, C) = (B, C) = 0$. A rigid movement keeps scalar products unchanged, so A, B, C can only go to points A^*, B^*, C^* for which $(A^*, A^*) = (B^*, B^*) = (C^*, C^*) = 1$ and $(A^*, B^*) = (A^*, C^*) = (B^*, C^*) = 0$. This means that OA^*, OB^*, OC^* must also be unit vectors and mutually perpendicular – exactly what we would expect on geometrical grounds.

Suppose we have found three points A^*, B^*, C^* that satisfy these conditions. The linear transformation determined by $A \to A^*$, $B \to B^*$, $C \to C^*$ makes $xA+yB+zC \to xA^*+yB^*+zC^*$. Will this linear transformation be a rigid movement? On page

181 we saw that a transformation was certain to be a rigid movement if it kept all scalar products unchanged in value.

Let us see what happens to the scalar product (P, Q) where $P = x_1A+y_1B+z_1C = (x_1, y_1, z_1)$ and $Q = x_2A+y_2B+z_2C = (x_2, y_2, z_2)$. The calculation of (P^*, Q^*) is a good example of how we benefit from our knowledge of the algebraic properties of the scalar product. We have $(P^*, Q^*) = (x_1A^*+y_1B^*+z_1C^*, x_2A^*+y_2B^*+z_2C^*)$. This, when fully multiplied out, contains nine terms. It begins $x_1x_2(A^*, A^*)+x_1y_2(A^*, B^*)+\ldots$. But now we see that the work is not going to be heavy, for $(A^*, A^*) = 1$ and $(A^*, B^*) = 0$. In fact, of these nine terms, six are simply nought because of the conditions satisfied by A^*, B^*, and C^*. In the remaining three terms, we use the fact that $(A^*, A^*) = (B^*, B^*) = (C^*, C^*) = 1$ and the whole expression reduces to $x_1x_2+y_1y_2+z_1z_2$. But this equals the original scalar product, (P, Q). So the linear transformation does in fact preserve all scalar products and hence is a rigid movement about the origin.

It is not difficult to round the whole thing off by proving that *every* rigid movement is in fact a linear transformation, but we will not go into this.

The arguments we have used are applicable to five, or any finite number, of dimensions. In five dimensions we would merely have the extra labour of writing A, B, C, D, E instead of A, B, C.

Scalar products also play a prominent part in the theory of Hilbert space, which is a generalization of Euclidean geometry to an infinite number of dimensions.

POSTSCRIPT ON LOGIC

A certain logical trap may perhaps lie concealed in this chapter, and should be pointed out. We will discuss this in terms of two dimensions; it is not essentially different in n dimensions.

In this chapter, points have been specified as (x, y) and we have drawn freely on the idea of Chapters One and Two. We have added vectors and multiplied them by numbers. In this sense we can speak of the line segment OA as consisting of all the points tA with $0 \leqslant t \leqslant 1$. As A is $(1, 0)$ this means all the points $(t, 0)$.

But this chapter gives us another way of defining the line seg-

ment OA, for Definition Two tells us what is meant by distance, and we could define the line segment as *the shortest route from O to A*.

Will these two definitions agree? In fact, as our geometrical experience leads us to expect, they do. This, however, ought to be discussed and proved. The agreement is due to the particular formula used in Definition Two; with other formulas it might not happen. For example, imagine an ordinary piece of squared graph paper, with coordinates x, y in the usual manner. We have now complied with Definition One; each point is specified by a pair of numbers (x, y). Now suppose that Mercator's map of the world is printed on this graph paper, and that we define the distance between any two points on the paper by giving the actual distance on the earth's surface between the places that these points represent. Suppose L and D are the points on the map that represent London and Delhi. The line-segment LD, as defined in Chapters One and Two with the help of the expression $(1-t)L+tD$, would then be the straight path that we would get by putting a ruler on the paper and joining L to D. But an aeroplane flying from London to Delhi by the shortest route would follow a very different path; on the map it would appear curved. Thus the vector approach and the distance approach would lead to different definitions of straight line.

We will now show that this complication does not arise with Definition Two. In particular, we will show that the shortest path from O to A agrees with the line segment already specified by the vector approach.

To select from all the smooth curves that join O to A the one that has the least length is a perfectly good mathematical problem. However, its solution involves the calculus of variations, which is not widely known. We can get round this difficulty in the following way. If P is not on the direct route from O to A, we would expect the distances OP and PA to add up to more than OA. If, however, Q is on the direct route, we should expect to find $OQ+QA = OA$. In Figure 80, O is the origin, A is $(1, 0)$, Q is $(t, 0)$, and P is (t, h). We assume $0<t<1$. By using Definition Two we can calculate the distances OA, OQ, QA, OP, PA. We find, of course, $OQ = t$ and $QA = 1-t$ while $OA = 1$. Thus $OQ+QA = OA$ and

this confirms our belief that Q is indeed on the direct route from O to A. It is otherwise with P. For $OP = \sqrt{t^2+h^2}$ which is clearly bigger than t, and $PA = \sqrt{(1-t)^2+h^2}$ which is clearly bigger than $1-t$. (We are assuming, as implied by Figure 80, that $h \neq 0$.) So $OP+PA$ is certainly more than $t+(1-t)$, that is, more than 1.

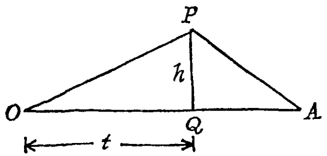


Figure 80

This agrees with our belief that P is not on the shortest route from O to A.

This indicates that we can produce a definition of direct route, *based on the idea of distance*, that leads us to the same concept of line-segment as we had from vectors.

Of course we ought to prove this result for any two points and not only for the specially simple points O and A.

Quite an interesting little exercise in algebra is to try to prove it directly for any two points H, K. What we have to show is that, unless P is of the form $(1-t)H+tK$, with $0 \leqslant t \leqslant 1$, the sum $HP+PK$ is bigger than HK.

An easier way out is to use the fact that distances are unaltered by translations and rotations. We first show, much as we did for OA, that any piece of the x-axis is a shortest route. If we apply any rotation, and then any translation, to a shortest route, this will necessarily produce another instance of a shortest route. So the shortest route definition of a line is equivalent to saying 'a line is anything that can be obtained from the x-axis by rotation and translation'. It is then easy to verify that this agrees with the definition of line used in Chapters One and Two.

CHAPTER TEN

Metric and Banach Spaces

MATHEMATICS often takes a metaphor and turns it into a tool. Some picturesque phrase of everyday life is taken literally and is shown to be both precise and logical. This has happened with the metaphor of *distance* which runs through much of our thinking. We say that a firm is *close* to bankruptcy. Unsuccessful generals often believe their strategy and tactics to be *close* to Napoleon's. A rumour may be *a long way* from the truth. We will be understood if we assert that Handel and Haydn are close to each other, but a long way from Indian music on the one hand or rock-'n-roll on the other. It would be an interesting problem in operational research to determine just what factors – harmony, rhythm, etc. – make us feel two musicians to be near together or widely separated. An immense chart might be produced, showing all musicians in their natural groupings – a kind of geometrical classification. One could even prove simple theorems – Haydn is close to Handel; Handel is a long way from Indian music; therefore Haydn is a long way from Indian music (see Figure 81).

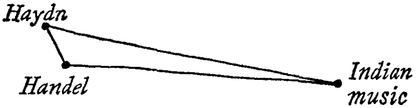


Figure 81

In all these illustrations, the distances have been between objects of considerable complexity – the condition of a business, a way of waging war, the style of a composer. These objects appear as *points*. In Figure 81, there is a point that represents Handel. In the mathematical theory of metric spaces also, each point may stand for some complex object – though nothing half as complicated as a musician or a general. A point may denote a vector, a

matrix, a curve, a function, or an operation (such as integration). We shall find precise ways of defining the distance between, say, one matrix and another, and similarly with the other classes of objects.

A *metric space* is a collection of objects for which some reasonable definition of distance has been made. We will explain later just what 'reasonable' means.

There may be several different, yet sensible, ways of arranging the same collection of objects. We have considered arranging musicians in accordance with their style of composition. A commercially minded person might want to arrange musicians in the order of their financial success; the musicians would then be arranged in a line, those who made most money being at one end, those who made least at the other. Someone else might be interested in the geographical distribution of musicians, and would arrange musicians according to place of birth. He would visualize the musicians arranged on the surface of the globe.

These three arrangements would count as different metric spaces, in spite of the fact that they have all been manufactured from the same material, namely, musicians. The geometry of objects arranged in line is clearly different from that of the same objects on a globe. Indeed in mathematics we are mainly concerned with the *arrangement*, the shape the objects form; we are more or less indifferent to what the individual objects are. If we were doing some work in which it was frequently necessary to refer to these three arrangements, we might use special abbreviations for them. The first might perhaps be called $S(M,c)$; here S indicates that we are talking about a *space*, M that its points represent *musicians*, and c that distances are defined in terms of style of *composition*. The second arrangement, with the musicians in line, could be called $S(M,f)$ since the distances are based on *finance*. The third might be $S(M,g)$ with distances determined by *geography*.

The three arrangements of musicians correspond to three different interests or purposes. In mathematics also, we may use different definitions of distance for the same collection of objects, but for different purposes. In Figure 82, in each of diagrams A, B, C we see a pair of curves $y = f(x)$ and $y = g(x)$. In each case we ask – should the curves be regarded as close together or not? Is

$y = g(x)$ a good approximation to $y = f(x)$? In each case, in certain circumstances, the answer could be yes. In A, $y = g(x)$ goes a long way from $y = f(x)$, but it only stays away for a very short time. If we were mainly interested in the areas under the two curves, it might well be that these areas would differ by very little, so the curves shown in A could be regarded as close together. So too, with this criterion, would be the pairs of curves in B and C. However, it might be that we want to make tables – such as those correct to a specified number of places – in which we could guarantee that no single entry was in error by more than a stated small amount. Situation A, in which there is a large, even though localized error, is no longer acceptable. By this test, the curves in A are not close together; those in B and C still are. But there are purposes for which the curves in C ought not to be regarded as close together. These curves behave very differently. The curve $y = g(x)$ has many wobbles in it; it represents a very much longer route than the other. In an investigation where we were particularly concerned to discover whether wobbles occurred or not, or where the length of the curve was important to us, we might regard the curves in C as widely different. We could meet this situation by saying that we would regard two smooth curves as close together only if they agreed approximately both in position and direction. This is not the case in C; where the curves cross, their tangents make quite a large angle. The curves in B would still qualify as close together.

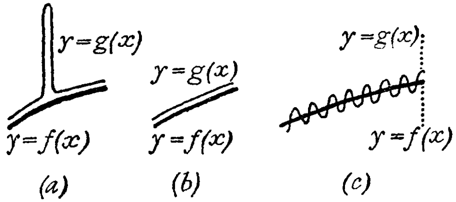


Figure 82

QUALIFICATIONS FOR DISTANCE

Not every situation has a simple geometrical representation. For example, in social life it is quite possible for Brown to be very friendly with both Smith and Jones, but Smith and Jones hate each other. It is impossible to show this situation by a diagram in which friendship is indicated by closeness. For Smith would have to appear close to his friend Brown, and Jones too should be close to Brown, but this makes Smith close to Jones, which does not correspond to the facts. This difficulty has indeed bedevilled both personal relationships and international diplomacy since the beginning of recorded history.

Accordingly, it is desirable to have some way of testing whether a situation can be visualized geometrically in terms of distance or not. We are led to analyse the meaning of distance; what properties do we assume when we speak or think in terms of distance? We measure distances by numbers; it is 3 miles from A to B. The number is never negative; we never say that a place is -5 miles away. The number can be 0 but only in a very special case – the distance from a place to itself. Distances are the same either way; the distance from London to Cambridge is the same as the distance from Cambridge to London. Finally, we cannot shorten a journey by breaking it. If we go from A to C and then from C to B, we must have gone at least as far as from A to B. This can also be looked at the other way round – if we walk 10 miles and then 4 miles, we cannot end up more than 14 miles from our starting point.

Set out formally, these requirements for *distance* are as follows:

1. For any pair of objects A, B, the distance from A to B is defined. We denote it by $d(A, B)$;
2. $d(A, B)$ is a real number and never negative;
3. $d(A, B) = 0$ when, and only when, A and B coincide;
4. $d(A, B) = d(B, A)$;
5. $d(A, C)+d(C, B) \geqslant d(A, B)$.

The theory of distance is concerned with the consequences of these axioms. The properties are very simple and there are very few of them, so distance geometry is not difficult.

It can be verified that all these properties hold for Euclid's distance, as given in Definition Two of Chapter Nine. The verification of property 5 calls for some skill. The other properties are easy to check.

Of course, the information contained in the five properties above is much less than that in the axioms of Euclidean geometry. When we imagine a metric space, we should not assume that it necessarily resembles Euclid's geometry. The five properties hold for distances on the surface of the Earth, which is quite unlike a Euclidean plane. The properties hold for distance as experienced by a caterpillar crawling on the surface of a statue or on a stationary bicycle, or a man exploring a maze; in such cases distance is always understood to denote the distance by the most economical route. It does not matter if there are several economical routes – as, on the Earth, from the North to the South Pole.

One can construct examples of metric spaces by defining $d(A, B)$ as the time it takes to get from A to B. The circumstances have to be suitable – for instance, on a hillside, it might take longer to climb from A to B than to descend from B to A, and property (4) would fail. It might not apply to road travel, since traffic jams would be worse at one time than another. Given freedom from such obstructions, one could construct a geometry, the points of which were the towns of Britain, and the distances the time required to drive from one town to another. The opening of a new motor road would warp this geometry in an interesting way.

A metric space can be defined by means of the king's move in chess. A king can move to an adjacent square in any direction. We define the distance between two squares as the smallest number of moves in which a king can get from one to the other. In Figure 83, the white king needs four moves to reach the square of the black pawn. The distance between the squares is therefore 4. Note that the pawn is four steps across and two steps up from the king, and that the larger of the numbers just mentioned gives the distance. For the king must use four moves in order to take four steps across; the smaller number 2 causes no difficulty, since it costs the king nothing to move upwards as well on any two of his four moves.

In this space a circle is a square. In any two-dimensional metric space, the circle, centre A, radius r, is defined as consisting of all

the points P at a distance r from A; that is, all P such that $d(A, P) = r$. In Figure 84, the black dots form a circle, centre the white king, radius 2.

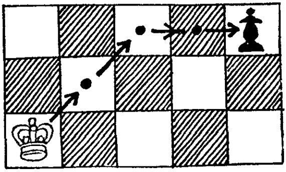


Figure 83

It is perhaps surprising that the Chess King's Metric, just described, should be mathematically significant. It leads directly to a metric space frequently used in mathematics and capable of wide generalization. If we imagine a chessboard which was divided not into sixty-four squares but into, say, a million, we could still define distance by the number of moves a king needed. The board has been so finely subdivided that we are finding it hard to distin-

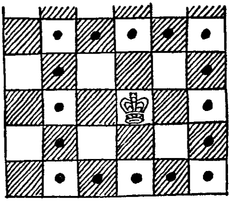


Figure 84

guish from a continuous plane. This suggests that we adapt the rule for finding distances so that it becomes meaningful for the plane. Given any two points, consider the difference of their x coordinates and the difference of their y coordinates. The larger

of these numbers defines the distance. In symbols,* the distance between (x_1, y_1) and (x_2, y_2) is defined as the larger of the numbers $|x_2 - x_1|$ and $|y_2 - y_1|$. In this space also, as we would expect, a circle is a square.

BANACH SPACES

Recent mathematical publications make frequent reference to Banach spaces. A layman, or a mathematician with a traditional background, might well wonder what these are. Has Banach discovered some special new shapes? I remember how illuminating I found it when, having heard Banach spaces mentioned, I looked up Banach's original paper to see what it was all about.† Banach begins by pointing to about ten situations in traditional mathematics (calculus) to each of which essentially the same argument applies. He argues, in effect, 'Why have the same proof ten times in ten different books? Why not prove it in one book, once and for all, and let the other nine refer to that proof?' His actual words (translated) are:

> This present work has the object of establishing certain theorems that hold in several different branches of mathematics, which will be specified later. However, in order to avoid proving these theorems for each branch individually, which would be very wearisome, I have chosen a different way, which is this: I consider in a general way sets of elements for which I postulate certain properties. From these I deduce theorems and then I prove for each separate branch of mathematics that the postulates adopted are true of it.

Banach spaces, then, are not something new. We have all worked with them but, like Molière's character who had been speaking prose all his life, we have not been aware of it. Sceptically minded

* The symbol $|x_2 - x_1|$ used in this definition is itself a measure of distance in the real line. It tells us how far x_2 is from x_1; we do not care whether x_2 lies to the right or left of x_1; The same symbol when applied to complex numbers denotes modulus or absolute value; it is again a distance measure, for $|z_2 - z_1|$ tells us how far z_2 is from z_1 in the Argand diagram.

† S. Banach, 'Sur les opérations dans les ensembles abstraits et leur application aux équations intégrals', *Fundamenta Mathematica*, 3 (1922), pages 133–81.

readers may feel inclined to question the statement 'we have all worked with Banach spaces'. I would point out that one very simple example of a Banach space is – a straight line.

THE NEVER-NEVER PRINCIPLE IN MATHEMATICS

To arrive at the idea of a Banach space it seems wise to begin with a familiar theme and show how, by generalizing it, we can reproduce in a miniature way the thinking of Banach.

In classical mathematics there are many numbers and functions which are defined but which we should be unable to produce if challenged to do so. The number π is an example. It could be defined as the area inside a circle of unit radius (in Euclid's geometry). But we should be in difficulty if someone asked, 'What number *exactly* is that?' We can give approximations to it, such as $\frac{22}{7}$ or $\frac{355}{113}$, and in fact we can produce approximations as close as anybody may demand. But the number itself is irrational and cannot be exhibited.

We met another example in Chapter Six. It is impossible to give the exact value of $\frac{1}{2}(1+\sqrt{5})$, but the Fibonacci sequence provides the approximations $1/1, 2/1, 3/2, 5/3, 8/5, 13/8 \ldots$ from which $\frac{1}{2}(1+\sqrt{5})$ can be calculated to any desired degree of accuracy.

An infinite series shows the same effect. If we define e by the series $1+(1/1!)+(1/2!)+(1/3!)+\ldots$, what we mean is that the sum of the first n terms of this series gives an approximation to e and that the approximation can be made as good as you like by taking n large enough.

Not only numbers but also functions are specified in this way. Infinite series are sometimes used as a convenient way of computing some function that could be defined explicitly; thus the infinite series $1+x+x^2+x^3+\ldots$ is sometimes easier to handle than the explicit form $1/(1-x)$. But the majority of functions used in analysis cannot be given explicitly; they can only be specified as a limit that is approached but never reached. This is true even of such well-known functions as $\sin x$ and $\cos x$; e^x can be specified by an infinite series or as the limit of $\{1+(x/n)\}^n$; $\sin^{-1} x$ can be defined by an infinite series or as the integral $\int_0^x (1-t^2)^{-\frac{1}{2}}dt$,

and this integral in turn is defined only as the limit of an unending process. These examples use functions familiar from elementary work. A function can be introduced simply by giving a series; we might say, let us investigate the properties of the function defined by $x+x^4+x^9+x^{16}+x^{25}+\ldots$, where the indices are the square numbers. This arises in the theory of elliptic functions.

Some mathematicians and mathematical philosophers have taken the extreme viewpoint that such numbers as π may not be used in mathematics at all, since their definition supposes an infinite process to have been completed, which is impossible. Whatever the logic of the position, the great majority of mathematicians, and of non-mathematicians too, make free use of such definitions. Mathematics would be very much constricted if these definitions were forbidden.

However, since about A.D. 1760 it has increasingly been realized that infinite processes must be used with very great care. All kinds of paradoxes can arise if we assume that some sequence or series defines a number when in fact it does not. In the eighteenth century great liberties were taken with series. For instance, the series $1+x+x^2+x^3+\ldots$ approaches $1/(1-x)$ provided x lies between -1 and $+1$. Some eighteenth-century mathematicians, in the belief that mathematical patterns will always look after themselves, would cheerfully put $x = 3$ and prove some result by reasoning which assumed that $1-3+9-27+81-\ldots$ was a valid expression for $-\frac{1}{2}$. Fourier, in section 218 of his epoch-making book *Analytical Theory of Heat* (1822), based a proof on the assumption that $1-1+1-1+1-1+\ldots.$ meant $\frac{1}{2}$.

The great mathematician N. H. Abel, in a letter written in 1826, described such practices as 'diabolical'.* Another critic of this loose reasoning was Bernard Bolzano (1781–1848), professor of philosophy and religion at Prague. He pointed out that, if you accepted the series $S = 1-1+1-1+1-1\ldots$, you could prove S to be anything you liked. By writing $S = (1-1)+(1-1)+(1-1)+\ldots$ you show $S = 0$; by writing $S = 1+(-1+1)+(-1+1)+\ldots$, you show $S = 1$; by writing $S = 1-(1-1+1-1+1-\ldots) = 1-S$, you show

* According to Oystein Ore. In Abel's Collected Works the milder phrase 'really fatal' appears. I do not know which version Abel actually wrote.

$S = \frac{1}{2}$. It is clearly most unsatisfactory to base mathematical proofs on a procedure capable of demonstrating $0 = 1 = \frac{1}{2}$.

We are thus led to distinguish between well-behaved or *convergent* series, such as $1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots$, and badly-behaved or *divergent* series,* such as $1 - 1 + 1 - 1 + \ldots$. There are even finer shades of distinction, between very well-behaved (or absolutely convergent) series, to which almost any reasonable process may be applied with confidence, and fairly well-behaved (or conditionally convergent) series, which have to be handled with some caution. We are going to look at a rather simple result, concerned with very well-behaved series. It was known in the nineteenth century for real and complex numbers. Banach showed that it could be extended to situations of a much more general kind.

THE CHAIN PICTURE

In order to develop this theme, it will be convenient to picture a series as a chain. In Figure 85, we see part of a chain representing the series $1 + \frac{1}{2} + \frac{1}{4} + \ldots$. The first link is 1 foot long, the second $\frac{1}{2}$

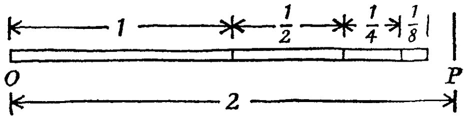


Figure 85

foot, and so on. However many links of this chain we may forge, we shall never pass the point P, 2 feet from O, though we may approach as close to P as we like. The convergence of the series is shown by this approach of the chain towards the point P.

I throw out a question which will not be answered at present – the answer is implicit in our later work. Suppose I prescribe the numbers $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}$, etc., as the terms of a series, but leave the choice

* A very readable historical account of divergent series, and the way in which divergent series can sometimes be rescued, can be found in Chapter Thirteen of K. Knopp, *Theory and Application of Infinite Series*, 2nd ed., Hafner Publishing Co., New York, 1948, and Blackie, 1951.

of signs to you. You are to insert + and − signs at will, so that you could write, for example, $1-\frac{1}{2}-\frac{1}{4}+\frac{1}{8}-\ldots$. Could you, by maliciously choosing the pattern of + and − signs, produce in this way a divergent series?

We now pass to two dimensions. If we consider the series $1+x+x^2+x^3+\ldots$ with $x=\frac{1}{2}i$, where $i=\sqrt{-1}$, we obtain the series of complex numbers

$$1+(i/2)-(1/4)-(i/8)+(1/16)+(i/32)-\ldots.$$

The chain now appears in the Argand diagram as in Figure 86. It spirals round and round, and is evidently approaching a point

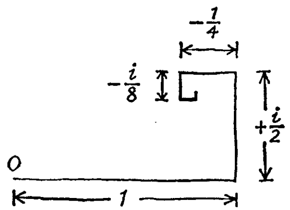


Figure 86

which, in fact, corresponds to the value we hope for, namely $1/(1-\frac{1}{2}i)$.

Figures 85 and 86 show the same chain in two different positions. In each position the chain approaches a point, that is, it represents a converging series. One naturally wonders; would it be possible to place this chain in such a way that it did *not* represent a convergent series? If I fastened the beginning of the chain to the origin, and let you choose the direction of each link, would it be possible for you to choose these directions in such a way that the chain perpetually wandered about, without ever settling down and approaching some point?

Before you have made any of your choices at all, is there anything I can predict about the course of your operations? Or can you go anywhere you like with this chain? It is fairly obvious that there is a restriction on how far you can go. However many links you use, their total length never reaches 2. You are predestined never to escape from circle I of Figure 87, with centre O and radius 2.

Now suppose you tell me you have decided to put the first link in the position OA. Can I now improve my prediction? Yes; the chain remaining at your disposal cannot take you a distance from

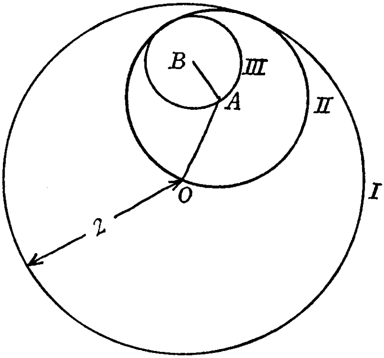


Figure 87

A exceeding 1; your prison is now circle II, centre A, radius 1. If you choose AB for the position of your second link, your prison shrinks to circle III, centre B, radius $\frac{1}{2}$. And so it continues; each decision you make halves the radius of the circle containing all your future possible positions. The walls of the prison continually close in, squeezing you down to the neighbourhood of some point. Which point it shall be is within your control; you can arrange things so that the chain approaches any point at all inside or on the circle I. What you cannot do is to escape settling down somewhere; you cannot be an eternal commuter – not with this chain.

We have used the particular series $1 + \frac{1}{2} + \frac{1}{4} + \ldots$ for the lengths of the links in the chain, but the argument could easily be adapted to other chains. The essential point is that, although we may use as many links as we like, the length of the chain can never exceed a fixed amount (the radius of circle I). *When the length of the chain is limited in this way, the chain is bound to approach some point.* This result appears in the classical theory of complex numbers in

the form of an apparently foolish theorem: *if a series is absolutely convergent, it is convergent.*

This theorem is not so foolish as it sounds. In it *convergent* means that the chain approaches some point; *absolutely convergent* means that, as we take more and more links, the *length* of the chain approaches some fixed, finite number. It is most helpful to know that the second of these implies the first, for the length of the chain is given by a series of numbers, while the wanderings of the chain itself are given by a series of vectors. In fact one could make a very tricky little examination question on this theorem, by giving a complicated way of choosing the directions of the links of a chain. The chain could be the one we have already used, with links of 1, $\frac{1}{2}$, $\frac{1}{4}$, etc. The first link could make an angle of 1° with the x-axis, the second 4°, the third 9°, and generally the n^{th} an angle of n^2 degrees. The student would be asked whether this prescription gave a convergent series of vectors, i.e. whether the chain approached some point. The student would find himself in great difficulties if he tried to calculate the position of the end of the n^{th} link. The point of the question is that he does not need to do so; the length of the chain is given by the convergent series $1+\frac{1}{2}+\frac{1}{4}+\ldots$, so the chain itself is bound to settle down near some point.

In the Argand diagram, any complex number z is represented by a vector in the plane. The length of this vector is written $|z|$, and called the modulus or absolute value of z.

The chain itself (as in Figure 86) corresponds to the series of complex numbers (or vectors):

$$S = z_1+z_2+z_3+\ldots.$$

The length of the chain is given by the series of real numbers:

$$L = |z_1|+|z_2|+|z_3|+\ldots.$$

Our theorem is that, if the series for L converges, we can be sure the series for S converges.

GENERALIZING THIS RESULT

This result is a useful one and is capable of considerable generalization. It is evidently not restricted to two dimensions. If we were

given the same chain as in our earlier example, with its beginning attached to a point O in the middle of a large room, and we were allowed to arrange the links anywhere within that room (i.e. not merely in a plane), we should still find that, whatever we did, the chain would always approach some point. The argument would be exactly the same except that, in Figure 87, we would have to speak, not of the circles I, II, III but of the spheres I, II, III. We could extend the result to spaces of even higher dimension without difficulty.

To make the result available for as wide a range of applications as possible, we examine what properties are being used in our theorem and its proof. The series $S = z_1 + z_2 + z_3 + \ldots$ involves the sign for *addition*; this will be meaningful if we are dealing with a *vector space*. The series for L involves the lengths of the links in the chain. The length of a link is the *distance* between its ends; we must suppose that distance is defined, that is, that we are dealing with a *metric space*. A third requirement is not so obvious. Imagine someone who is familiar with complex numbers but has never met irrational numbers. For such a person the complex plane consists of all the numbers $x+yi$ with rational numbers x, y. Now it is possible to arrange the chain so that it approaches any point inside the circle I. Suppose then we prescribe the chain so that it approaches the point $(1+i)/\sqrt{2}$. This point lies inside the circle I. We can approach it by using a chain prescribed entirely by means of rational numbers. The chain is thus recognized and accepted by the person in question. He can show, just as we have, that there are circles I, II, III . . . closing in, forming smaller and smaller prisons. We know that there is just one point within all these circles, the point $(1+i)/\sqrt{2}$. But our person does not recognize that point as existing. So although the circles seem to be closing in, most satisfactorily, and pin-pointing a position in the plane, according to that person, they do not catch anything; when the net is hauled up, it is empty. It is because we find this situation very unsatisfactory that we are led to discuss irrational numbers.

If we regard the plane as consisting of all the points with real number coordinates x, y, then we can be sure that a system of circles, closing in, will indeed catch a point. Such a space is called *complete*. On the other hand, a space where circles can close in,

each circle lying within the previous one, and yet not catch any point, is said to be *incomplete*.

The three requirements are therefore that the space must be (1) a vector space, (2) a metric space, (3) a complete space. A space that meets these requirements is called a Banach space. Later on, the axioms for a Banach space will be stated in a more detailed form. For the moment, we simply want to indicate the direction of our train of thought; we are looking for situations of the widest possible variety in which we may hope to prove a theorem about absolute convergence implying convergence.

We must explain how we propose to generalize the idea of *circle* or *sphere*. No originality or great mental power is needed to do this. There is, however, a little terminology that can usefully be introduced. In traditional geometry, the word *circle* is used a little ambiguously. Sometimes it refers to the points at a distance r from a centre O, and these only, as when we say 'the locus is a circle'. Sometimes, as when we speak of 'the area of a circle', we mean the curve just mentioned, together with the region inside it. In order to have a quick way of specifying exactly what is meant, the following three definitions have been devised. These hold in any metric space, i.e. in any system where distance is defined. The sphere, centre A, radius r, consists of all the points P at a distance r from A, i.e. all points P such that $d(A, P) = r$. The *open ball*, centre A, radius r, consists of all the points inside this sphere, that is, all the points whose distance from A is less than r; $d(A, P) < r$. The *closed ball* consists of all the points that are either on the sphere or inside it, that is, all the points P with $d(A, P) \leqslant r$. The sphere is the skin of the closed ball; the open ball is what remains when (as with an old cricket ball) the skin peels off.

Note that in two dimensions, a 'sphere' means a circle. In one dimension, a 'sphere' means a pair of points. These may sound a little odd at first; we cannot avoid this oddness if we are to have a uniform terminology for any number of dimensions, finite or infinite.

In our earlier work with the circles I, II, III ... our result would have to be stated in terms of closed balls. Consider the simplest possible example, the series $1 + \frac{1}{2} + \frac{1}{4} + \ldots$ as illustrated in the Argand diagram. The circles I, II, III ... are then as shown in

Figure 88. These circles all touch at the point P representing the number 2, and P belongs to all the closed balls of the system; it is

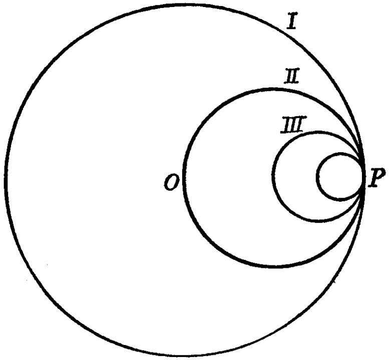


Figure 88

'inside or on' all the circles. But there is no point belonging to all the open balls of the system; no point is actually *inside* all the circles.

AN INTEGRAL EQUATION

The title of Banach's paper mentioned 'applications to integral equations'. Accordingly, we consider an extremely simple integral equation. If we integrate e^x between the limits 0 and x we get $e^x - 1$. Let us write simply f for the function e^x and simply $\int$ for the operation of integration from *0* to x. We obtain the equation:

$$\int f = f - 1. \tag{1}$$

Now this equation is characteristic of e^x. No other function satisfies it. An enterprising sixth-form teacher might decide to introduce e^x to his pupils as the solution of this equation. It would probably not be a wise thing to do but mathematically it would be sound. What do you think would be the reaction of the class if the teacher proceeded as follows?

Equation (1) may be rewritten

$$1 = f - \int f. \tag{2}$$

Certainly no objection would be raised to this step. Equation (2) may be written

$$1 = (1 - \int)f. \tag{3}$$

There might be objections to this, but they could be met. Equation (3) is merely a linguistic change: the use of symbols in it is defined so that it merely reaffirms equation (2).

Divide both sides by $(1 - \int)$.

$$f = \frac{1}{1-\int} \cdot 1. \tag{4}$$

Expand the fraction as the geometrical progression with common ratio $\int$.

$$f = (1 + \int + \int\int + \int\int\int + \ldots)\ 1 \tag{5}$$

Now carry out the integrations indicated. Integrating 1 from 0 to x gives x. Applying the operation a second times gives $x^2/2$. The third application gives $x^3/6$. Hence:

$$f = 1 + x + (x^2/2) + (x^3/6) + \ldots . \tag{6}$$

It is not hard to check that the general term of the series (6) is $x^n/(n!)$, as it should be in the series for e^x. So this procedure has at any rate the merit of yielding the correct answer.

I have tried this on many meetings of sixth-formers and the calculations above have invariably been greeted with laughter. The joke, however, is that these calculations are perfectly rigorous, and represent a standard procedure in the solution of integral equations. I do not want to be responsible for a large number of failures in public examinations. In case the impression is created that anything whatever is allowable in mathematics, I would hasten to add that the above calculations are rigorous only when used by someone who quotes the theorems that justify each step being taken. In fact, to justify the work above we would have to prove three things – (i) that the series (6) converges, that is, that it

really represents something, (ii) that this something is actually a solution of equation (1), and (iii) that nothing else is a solution.

One can indeed prove that the series (6) converges however large x may be. We shall be rather conservative in our treatment of it; we shall prove only that it converges for $0 \leqslant x \leqslant \frac{1}{2}$. The reason for this is that the geometric progression $1 + \frac{1}{2} + \frac{1}{4} + \ldots$ has been a theme running through our work; we shall be able to prove our result by exactly the method we have already used, that is, by considering a chain with links whose lengths are the terms of this geometrical progression.

The chain will lie in a particular space, obtained as follows. Imagine we have an immense stock of labels, on each of which is the graph (or other specification) of a function continuous for x from 0 to $\frac{1}{2}$. We suppose that in heaven, or somewhere where there is a lot of room, these labels are arranged in an orderly way. As we saw at the end of Chapter Seven, functions continuous in the interval 0 to $\frac{1}{2}$ form a vector space. The labels must be arranged so as to bring this out. Also we want to talk about the lengths of links in a chain; we need to know what is meant by the distance between two labels. We want two labels to be close together if they represent functions that are nearly the same, far apart if the functions are widely different. A way of doing this is shown in Figure 89. This shows two graphs, $y = f(x)$ and $y = g(x)$. The arrow shows where the graphs are the greatest distance apart. This dis-

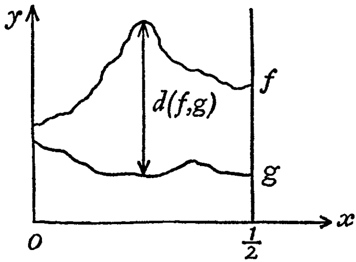


Figure 89

tance – the length of the arrow – is defined to be the distance between the two functions f and g; $d(f, g)$ for short.

It is not hard to see that this definition has the five qualifications for distance listed on page 190. For $d(f,g)$ is defined as a real number, never negative. If $d(f, g) = 0$, this means that the greatest distance between the graphs is 0, so they coincide; thus (3) holds. So does (4); $d(f, g) = d(g, f)$. Figure 90 indicates how we prove (5). We want to show that $d(f, g)+d(g, h)$ cannot be less than $d(f, h)$. Now $d(f, h)$ is the length PR, which can be broken into PQ and QR. But PQ cannot exceed DE, the maximum distance from f to g, and

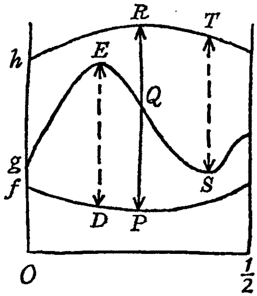


Figure 90

QR cannot exceed ST, the maximum distance from g to h. As $DE = d(f, g)$ and $ST = d(g, h)$, this is what we want. The argument still needs a little tidying up; for example, the objection could be raised that the graphs might not lie in the particularly simple way shown in Figure 90. This objection can be met without much difficulty.

So our space of functions has been shown to meet two of Banach's conditions; it is a vector space, and distance is defined in it. It remains only to show that it is complete, that when the spheres close in, they will catch the label for some continuous function. This also can be proved. Incidentally, in proving it we do not use any modern mathematics: we prove a certain result about continuous functions in exactly the way a nineteenth-century mathematician would have done.

Banach showed that, in any space satisfying his three conditions,

a chain was sure to converge if we knew that the total length of the chain was limited. So all we have to ask now is – what is the length of the chain corresponding to the series involved in equations (5) and (6) on page 203?

THE LENGTH OF A LINK

In our work with complex numbers, the length of the chain (see page 199) was given by the series:

$$L = |z_1| + |z_2| + |z_3| + \dots.$$

The terms of this series represent the lengths of the individual links in the chain. For complex numbers, $|z_1|$ measures the distance of the point z_1 from the point 0. This suggests that, in our work with functions, the length of the link f_1 should be the same as $d(f_1, 0)$, the distance of f_1 from the function 0. By the function 0 we mean the function with graph $y = 0$, for this is the zero vector (see page 153). To bring out the analogy with $|z_1|$, we denote the length of the link f_1 by $||f_1||$. We do the same of course with the links f_2, f_3 and so on.

For any continuous function f, then, by $||f||$ we understand

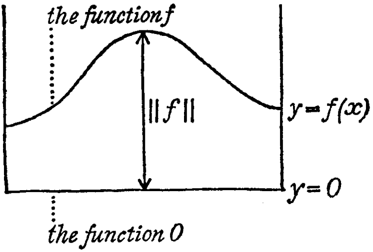


Figure 91

$d(f, 0)$. This is illustrated in Figure 91, where we see the two functions f and 0; the length of the arrow gives the greatest separation between the two graphs.

In algebraic language, $||f||$ is defined as the maximum value of $|f(x)|$. The absolute value sign, $|\quad|$, has to be used here, for $f(x)$ may take both positive and negative values. We are interested in the greatest distance that $f(x)$ can get from 0; we do not mind whether this happens above or below the line $y = 0$.

We shall refer to $||f||$ as *the size of the function f.* This is a new concept, not found in traditional calculus.

THE EFFECT OF INTEGRATION

Equations (5) and (6) contain series in which each term is found by integrating the term before. What is the effect of integration on the size of a function? Suppose we have some continuous function ϕ, with $||\phi|| = M$. This means, as shown in Figure 92, that the

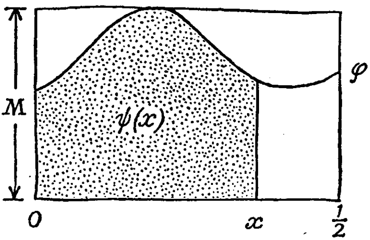


Figure 92

numerical value of ϕ reaches M, but never surpasses it in the interval we are considering. (Figure 92 has been drawn for the most easily visualized case, in which $\phi(x)$ is positive throughout. It can be shown without great difficulty that the conclusions to be reached hold equally well in the general case.) Let ψ denote $\int \phi$, the function obtained by integrating ϕ between 0 and x. Our aim is to estimate $||\psi||$, the size of ψ, given by the maximum value of $|\psi(x)|$. Now $\psi(x)$ can be visualized as the area under the graph of ϕ between 0 and x. In the figure, this area (shown shaded) lies entirely inside a rectangle of height M and base $\frac{1}{2}$. Accordingly, $\psi(x)$ cannot exceed $\frac{1}{2}M$. So, by our definition of $||\psi||$, it follows that $||\psi||$ cannot exceed $\frac{1}{2}M$. That is, the size of ψ is at most half the size of ϕ. Now

$\psi = \int \phi$, so we see that *integration halves the size of a function* – at worst: it may reduce it even more, and usually does.

Exercise

Calculate the sizes of the terms in series (6) on page 203. Find the ratio of each term to the preceding one, and verify that, with one exception, this ratio is actually *less* than $\frac{1}{2}$.

We have now reached our goal. The series in equations (5) and (6) correspond to a chain in which each link is half as long as the previous link, or less. Thus the series formed by the lengths of the links compares favourably with the geometrical progression of ratio $\frac{1}{2}$. However many links of the chain we may forge, the length of the chain can never exceed twice the length of the first link. We are therefore in a position to apply our idea of the contracting prison; we shall find a sequence of closed balls, each contained in the one before, and narrowing down to determine a single point, to which the chain itself approaches. We can be certain that the series (5) and (6) converge.

In a formal treatment the proof would not yet be complete. We would have to show, for instance, that the continuous function, to which the series converges, is in fact a solution of equation (1) on page 202. This gap we shall make no attempt here to fill. Our aim is simply to show the central idea, the power and the beauty of the analogy to which Banach called attention.

As was mentioned near the beginning of the chapter, there may be several different reasonable ways of defining the distance between two functions, depending on the purpose we have in mind. The definition of distance that we have been using is sometimes called the *uniform metric*. It is closely related to the (slightly troublesome) idea of *uniform convergence* in classical analysis. Where we simply say that the chain $f_1 + f_2 + f_3 \ldots$ is approaching the point S (by which we mean that the distance of the end of the chain from S is tending to nought), the classical analyst had to say that the series $f_1(x) + f_2(x) + f_3(x) + \ldots$ was *converging uniformly* to $S(x)$. The geometrical approach, outlined above, is a simple one to teach and is already being used in various places to replace the older approach to uniform convergence.

The space we have used, in which each point represents a continuous function, is sometimes called $\mathscr{C}(I)$. This sign indicates that we are dealing with continuous functions ($\mathscr{C}$ for continuous) on an interval, I. In our work the interval I was from 0 to $\frac{1}{2}$.

SERIES OF TRANSFORMATIONS

About two centuries ago, trigonometry became annexed to algebra through the discovery that all the elementary properties of the trigonometric functions were simply algebraic properties of $e^{i\theta}$. If we put $\theta = 1$, the expression reduces to e^i. Now i carries the interpretation 'turn through a right angle'. Thus, already in the mathematics of 1750, we find the remarkable instruction 'raise e to the power *turn through a right angle*'.

Now 'turn through a right angle' is a linear transformation of the type discussed in Chapter Three. This suggests that we may be able to attach a meaning to e^T, where T stands for any transformation, or, if you prefer, for the matrix representing the transformation. As e^x is most easily defined by a series, this suggests that we consider series involving the powers of a transformation. We shall certainly need to discuss whether the series converges or not. Our theorem about the convergence of a chain will come in very handy; we shall only be able to use it if we can find satisfactory definitions for $d(A, B)$, the distance from transformation A to transformation B, and $\| T \|$, the size of transformation T.

Now it is not unreasonable to suppose that we can define the distance between two transformations. Consider rotations as a particular example of transformations. Surely we can say that a rotation through 10° is pretty close to a rotation through 11°, but farther away from a rotation through 60°. Now rotations act on the points of a plane. We say that rotations A and B are close together if they have nearly the same effect on the points of the plane. Suppose A sends any point P to Q, and B sends P to R. We might be tempted to say that we shall regard A as close to B if, however P is chosen, the points Q and R lie close together. But this will not do. Suppose in fact that A and B were rotations through 10° and 11°, which we regard as close together. If P were chosen a million miles away from the origin, Q and R would be more than 16,000

miles apart. This we do not usually regard as a small distance. However, it is small in comparison with a million miles.

We could deal with the situation by saying that we would regard two operations A and B as having nearly the same effect if the distance QR was small in comparison with the distance OP; that is, we would use the ratio QR/OP as a measure of how differently the operations A and B affected the point P. Another way is to require that the distance OP should not exceed unity; then, if QR turns out to be large, this must be because A and B have widely different effects, and not because P is very far from the origin. We will adopt this second idea.

Accordingly, our procedure for finding out how much two linear transformations differ from each other will be the following. Let v stand for the vector OP; we require that the length of v must not exceed 1. The points Q and R are given by the vectors Av and Bv; the distance QR is the length of the vector $Bv - Av$. We choose v, subject to the restriction on its length, in such a way that the length of $Bv - Av$ is as large as possible. This maximum length of Bv—Av we then define to be $d(A, B)$, the distance between A and B.

This does in fact give us a very effective measure of the resemblance between the operations A and B. Suppose for example we had established for two particular operations, A and B, that $d(A, B)$ was 0·001. We would then know that, if we applied A and B to some vector v, of length 1 inch, the resulting points Av and Bv would be at most 0·001 inch apart. Suppose A was given by a complicated formula and B by a simple one, and that we were engaged in some task where an error of one part in a thousand did not matter. We could then replace the complicated A by the simple B without damage to our project.

Having defined the distance $d(A, B)$, we can now go on to define *the size of a transformation A*. As we saw in the section 'The length of a link', the size of a complex number z is its distance from the complex number 0, the size of a function f is its distance from the function 0. Naturally then, we explain the size of a transformation A as its distance from the transformation 0; that is, we define $\| A \|$ as $d(A, 0)$. Replacing B by 0 in our definition of $d(A, B)$, we find that $\| A \|$ means the maximum length that the vector Av can

have, subject to the condition that the length of v is not to exceed 1.

We have now reached something that is easily visualized. The condition that the length of v should not exceed 1 means that v must represent a point in or on the unit circle, $x^2+y^2 = 1$. To each such vector v the transformation A is applied, giving a new point Av. Suppose all the points Av that can be obtained in this way, marked on graph paper. They will fill a certain region which

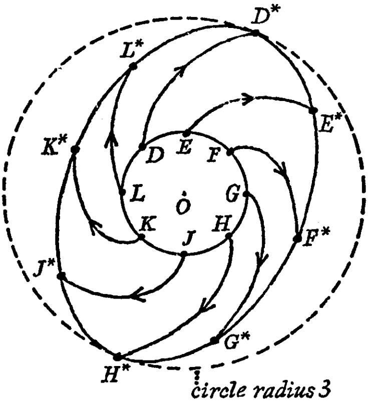


Figure 93

in fact is bounded by an ellipse. Figure 93 shows the unit circle and the ellipse to which it is sent by the transformation A; $x^* = x+2y$, $y^* = -2x+2y$. The points on this ellipse most distant from the origin are D^* and H^*. They are at distance 3. Thus if the length of v is 1 or less, the length of Av is 3 or less. The operation A never enlarges a vector more than three times. For the operation shown in this figure, $|| A || = 3$.

Figure 93 does not only enable us to visualize the meaning of $|| A ||$; it also helps us to see what $d(A, B)$, the distance between two operations, means. For $d(A, B)$ was defined as the maximum length of $Bv-Av$, on the understanding that the length of v was not to exceed 1. Now $Bv-Av$ may be written $(B-A)v$, and the maximum value of $(B-A)v$ is what we should consider if we were finding

$\| B-A \|$, the size of the difference $B-A$. So the distance $d(A, B)$ is the same thing as the size of $B-A$.

For example, suppose we wish to know the distance between the matrices $\begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$. The difference of these two matrices is $\begin{pmatrix} 1 & 2 \\ -2 & 2 \end{pmatrix}$, the matrix of the transformation illustrated in Figure 93. We have seen its size to be 3. Accordingly the distance between the two matrices is 3.

One may think of $\| A \|$ as the greatest magnification that the operation A produces in the length of any vector. If $\| A \| = a$ and $\| B \| = b$, we can see that $\| AB \|$ cannot exceed ab. For operation B magnifies at most b times and operation A at most a times; operation B followed by operation A cannot possibly magnify more than ab times. Usually the magnification will be less, for the choice of vector that gives maximum magnification at the first stage will not give maximum magnification at the second stage. This is still true even when $B = A$; as a rule $\| A^2 \|$ is less than $\| A \|^2$.

Exercise

The transformation A; $x^* = 3y$, $y^* = 2x$ has $\|A\| = 3$. The maximum magnification occurs for the vector (0, 1). Verify that $A^2 = 6I$, so $\|A^2\| = 6$, less than 3^2.

In the same way we can see that $\| A^n \|$ never exceeds, and is usually less than, $\| A \|^n$.

This section opened with the question whether it was legitimate to use the series e^A when A was a transformation or a matrix. We have seen that the size of A^n never exceeds a^n, where $a = \| A \|$. In the series for e^A then, each term has a size which is less than the corresponding term in the series for e^a. Now it is known that the series for e^a converges for every number a. So once again we are dealing with a chain of limited length, and the stage is set for an application of our earlier argument.

We began this section with rotations as an illustration; that is, our transformation A was a mapping of a plane to itself. But the argument would apply to a great variety of spaces. For example,

in the section 'The Effect of Integration' (page 207) we showed that, in the circumstances there considered, the operation $\int$ led to a function at most half the size of the original function. This allows us to attach a size to the operation $\int$; we have $\| \int \| = \frac{1}{2}$.

When we are dealing with transformations of a finite-dimensional space, such as a plane, every transformation A has a finite size $\| A \| = a$, and (after filling certain gaps in our argument) we can prove that, for every A, the series e^A is absolutely convergent. Note how very much simpler this argument is than going into the details of the numbers that would occur in the matrices for A^2, A^3, etc., found by the ordinary rules for matrix multiplication.

Caution is needed when dealing with transformations of infinite dimensional spaces; it is no longer true that every transformation A necessarily has a finite size $\| A \|$.

The brief sketch above suggests a great extension of the machinery at our disposal. We may start working with expressions such as $1/(1-\int)$ or $e^{\int}$ or $\sin \int$, subject always of course to proper precautions.

In Chapter Seven (see page 149) we saw that actuaries could, with justification, apply the binomial theorem to an expression such as $(1+\Delta)^n$. We now give another example of actuarial formalism, in which an infinite series in Δ is used with some freedom. We are dealing with an analytic function $f(x)$; we know its values only when x is a whole number, i.e. we have only $f(0), f(1), f(2) \ldots$; we want to estimate $f'(0)$, the slope of the function for $x = 0$. The argument runs as follows; by Taylor's Theorem,

$$\begin{aligned} f(x+a) &= f(x)+af'(x)+\tfrac{1}{2}a^2f''(x)+ \ldots \\ &= (1+aD+\tfrac{1}{2}a^2D^2+ \ldots) f(x). \end{aligned}$$

The series inside the bracket, of which we have written only the first three terms, is the series for e^{aD}. Thus

$$f(x+a) = e^{aD} f(x).$$

In this equation we put $a = 1$. The left-hand side becomes $f(x+1)$ which is $Ef(x)$ or $(1+\Delta) f(x)$. Omitting the symbol $f(x)$ on each side, we arrive at :

$$1+\Delta = e^D.$$

Taking logarithms we have:

$$D = \log_e(1+\triangle) = \triangle - \tfrac{1}{2}\triangle^2 + \tfrac{1}{3}\triangle^3 - \ldots.$$

The series used here is a standard one in the theory of logarithms but it is usually applied only to numbers. The result above claims to show how differentiation can be done by means of finite difference operations, which only involve whole numbers x.

The argument above has been given in a purely formal way. It would be an exercise for a student of functional analysis to turn this argument into a piece of real mathematics – to show just what assumptions must be made if we are to be certain that the result is correct.

THE AXIOMS OF BANACH SPACE

On pages 211–12, when we were trying to visualize the metric space of transformations, we proceeded in the following order; first, we managed to visualize $\| A \|$, the size of A; then we observed that $d(A, B)$, the distance between A and B, was the same thing as the size of the difference B–A. This order can be followed not only for visualizing, but also for defining distances. We can begin by defining the size, or length, of a single vector, and then define $d(A, B)$ as the length of $B-A$. This in fact is the order most commonly used in work with Banach spaces. The requirements for a Banach space may be stated as follows:

1. The system must be a vector space. It must pass the ten tests specified in the last section of Chapter Seven;

2. For every vector v, the size $\| v \|$ must be defined, in such a way that,

(1) $\| v \|$ is a real number, never negative,
(2) $\| v \| = 0$ when, and only when, $v = 0$,
(3) for any number k, $\| kv \| = | k | . \| v \|$,
(4) $\| u+v \|$ never exceeds $\| u \| + \| v \|$;

3. Then $d(u, v)$, the distance between u and v, is defined as $\| v-u \|$;

4. The space is complete; the contracting prisons, the closed balls, always contain some vector.

This list of requirements, as we saw earlier, was drawn up by Banach, because it covered several situations which had already played an important part in mathematics. In Chapter Four of the encyclopedic work, *Linear Operators*, by Dunford and Schwartz (John Wiley, 1958, 1963), a list is given of twenty-six Banach spaces – that is, twenty-six different topics in mathematics to which the above list of requirements applies. Some of these topics are rather advanced and will be recognized only by professional mathematicians; others, such as the examples given in this chapter, can be appreciated by someone with a good general knowledge of mathematics. Banach spaces therefore have plenty of applications. These applications lie in the middle and higher parts of mathematics. Banach spaces, almost certainly, would not be helpful to an engineering apprentice, concerned with the drawing of blueprints and the mathematics involved in cutting a piece of metal to the correct size and shape. Banach spaces, however, could be made of interest to a student working for an engineering degree involving calculus, infinite series, and matrices.

In a Banach space we have much of the machinery of Euclid's geometry; we have straight lines and planes, circles and spheres, parallelograms and bisected lines. But we are a little uneasy and unsure; we cannot clearly visualize the space; we are not certain whether these objects, with familiar names, will behave in the way we expect. The object of the theory of Banach spaces is to dispel this uncertainty. By deductive reasoning from the four axioms above, certain theorems are proved. We know we can rely on these. For example, in Banach space, are the opposite sides of a parallelogram of equal length? It is not hard to show that axiom 3, defining the distance $d(u, v)$, ensures that they will be. On the other hand, the theory warns us, by examples – such as the Chess King's Metric – of unexpected things that can happen in Banach spaces, such as circles being parallelograms.

A PARTICULAR METRIC SPACE

Around pages 209–10 we were looking for a way of defining the distance between two transformations. We were led to a definition by considering when two rotations could be regarded as close

together. We arrive at quite an interesting little metric space if we look into this matter more closely, and consider the movements about the origin discussed in Chapter Nine. These were rotations about the origin and reflections in lines through the origin.

Suppose we gave someone a lot of labels, specifying rotations about the origin, and asked him to arrange them in an orderly way that would bring out their relationships. He would naturally arrange them in order; 'Rotation of 0°', 'Rotation of 1°', 'Rotation of 2°' and so on. When he reached 'Rotation of 359°' it might strike him that rotation through 360° has the same effect as rotation through 0°, and that he ought to bring the two ends of the sequence close together. Probably he would place his 360 labels in a circle, and space them evenly. And this arrangement would in fact give the metric space for the rotations, in accordance

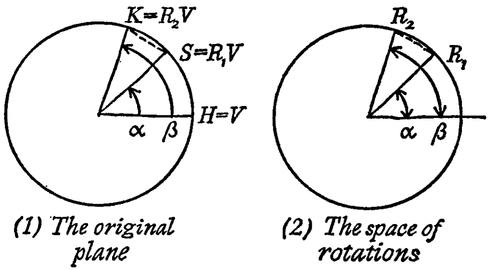


Figure 94

with the definition of distance stated on page 210, provided he chose a circle of radius 1. For suppose R_1 represents rotation through α and R_2 through β.

Figure 94 (2) shows the positions at which the labels would occur in the arrangement just described. The distance between these positions is the length of a chord of the unit circle, that subtends an angle $\beta-\alpha$ at the centre. According to the definition, $d(R_1, R_2)$ should be the maximum distance between R_1v and R_2v that can occur with a vector v of length not exceeding 1. Clearly this maximum will occur for a vector v whose length is actually 1. Figure 94 (1) shows H, corresponding to such a vector v, and S and

K, the points to which R_1 and R_2 carry H. It does not matter where H is chosen on the circle; the distance SK is always the same, and it is geometrically obvious that the length SK in (1) is the same as the distance from R_1 to R_2 in (2).

Now we come on to reflections. In Figure 95 we suppose that the

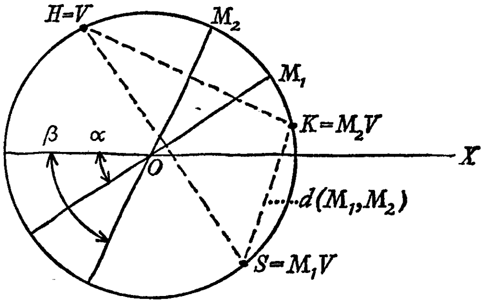


Figure 95

mirrors in which reflections M_1 and M_2 occur make angles α and β with the horizontal. Again we see H, a point on the unit circle, and the points S and K, to which M_1 and M_2 carry H. Once again it will appear that, wherever H is chosen on the circle, the distance SK is the same, and thus gives the required maximum separation. Suppose H has the position on the circle corresponding to the angle θ, i.e. OH is at an angle θ to OX. One can then show that S has the angular position $2\alpha - \theta$ and K the angular position $2\beta - \theta$. The difference between these is $2(\beta - \alpha)$, so this is the angle subtended by SK at O; in the metric space, the reflections M_1 and M_2 should be separated by a distance equal to SK. This will be achieved if we arrange the reflections round a circle of unit radius, with M_1 at angular position 2α and M_2 at 2β. The position of each reflection in the metric space is thus at an angle *twice* that which its mirror (in the original plane) makes with OX. The labels for the reflections will fill the unit circle and just fill it, for the angle between the mirror and OX varies from $0°$ to $180°$ only; the mirror

that makes 180° with OX specifies exactly the same reflection as the mirror at 0°. The factor *twice* just compensates for this.

So rotations fill a circle and reflections fill a circle. The interesting question is now – what is the distance between a rotation and a reflection? Is there any reflection which can be regarded as particularly close to a given rotation? The answer, as perhaps one might guess, is 'No'. Given any reflection M and any rotation R,

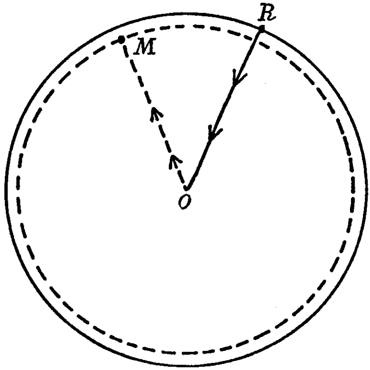


Figure 96

one can always find a vector v, corresponding to a point H on the unit circle, such that Mv and Rv lie at opposite ends of a diameter, and hence at a distance 2 from each other. This is clearly the maximum separation that can occur, so $d(M, R) = 2$. Any reflection is at a distance 2 from any rotation.

We can make a model of this metric space quite easily. In Figure 96, the circle shown by a continuous line is supposed to be on the front of the paper. It is of unit radius and the labels for the rotations are supposed to be arranged around it. The dotted circle is supposed also to be of unit radius, *and to be drawn on the back of the paper*. The labels for the reflections are arranged round it. The only way from the front of the paper to the back is through a little hole made in the paper at the point O. The shortest way from any rota-

tion to any reflection is thus to go straight in to O, pass through the hole, and go straight out again, as indicated by the path ROM. This distance is 2, as required.

NOTE ON FUNCTIONAL NOTATION

Nearly all recent books on mathematics indicate functions in a manner different from the traditional way. Why has the need for such a change been felt? Consider Figure 91 on page 206. Here we see a graph lying entirely above the x-axis. As was explained earlier, the size of such a function is given by the maximum value of $f(x)$. Now traditionally, we describe the function in question as being 'the function $f(x)$'. Accordingly, in traditional terminology, we would have to say, 'The size of the function $f(x)$ is given by the maximum value of $f(x)$.' Now this is most awkward; it seems to be saying that the size of something is the maximum value of itself – which is nonsense. And in fact '$f(x)$' is used in the sentence above in two completely different senses. The first time it is used it specifies a function, the second time a number. The function could be specified by drawing its graph; instead of defining the size of a function we could equally well have defined the size of a graph. Then, for a situation such as that in Figure 91, we could have said, 'The size of *a graph* is given by the maximum *height of a point on that graph above the x-axis.*'

In the traditional version of this, both the italicized parts of this sentence are replaced by the same symbol, $f(x)$. Yet they clearly signify different things – one the whole of the graph, the other the height of a single point on the graph.

Our thoughts hardly ever correspond to the literal meaning of the words we say; someone brought up in the traditional notation will not be disturbed by the use of the symbol in two different senses. His attention is concentrated, not on the particular words or symbols used, but on the actual problem in hand; the intention behind the sentence is clear to him, and he does not see what all the fuss is about.

One can certainly go too far in the direction of trying to get an extremely precise, consistent notation. In the U.S.A. I have certainly known students who found calculus much harder to learn

because it was presented in a 'modern' way, with long initial discussions of function, domain, and range. On the other hand, there are probably difficulties in the traditional approach to calculus, caused by confusing notation. We all know of students who romp away when drawing the graph $y = 2x+3$ but come to a sudden halt when asked to graph $y = 3$. The difficulty here is probably that they recognize $y = 2x+3$ as a way of specifying a function, but regard $y = 3$ as a way of specifying a number.

Function implies some kind of stimulus-response situation. With $y = 2x+3$, if you say 1 for x, I shall answer 5 for y; if you say 2, I shall answer 7, and so on. When drawing the graph of $y = 3$ we are envisaging the situation where, whatever you say, I shall answer 3.

One way of expressing this notationally, is to use the symbolism for a mapping. With $y = 2x+3$ we associate $1 \to 5$, $2 \to 7$ and so on; your stimulus $\to$ my response. We would thus speak of the function or mapping $x \to 2x+3$. When requiring the graph of $y = 3$, we would (on this approach) ask the pupil to illustrate the mapping $x \to 3$. The problem could be illustrated by an experiment carried out at constant temperature; if, x minutes after the start, you asked me the temperature, I should reply 3 degrees; the temperature would always be 3 degrees, whatever x. We had an example of such a constant mapping in Chapter Three, the bankruptcy mapping, $v \to 0$. Whatever your investment v, what you get out is the same – nothing.

With the system just described, we would not speak of the function x^2 but of $x \to x^2$; we would not speak of the function $f(x)$, but of $x \to f(x)$.

Another system starts from the idea that a function can be completely specified by its graph; it goes a step further and says a function *is* its graph. The graph of $y = x^2$ consists of the points (0, 0), (1, 1), (2, 4), (3, 9), $(\frac{1}{2}, \frac{1}{4})$ and many others. Now in the expression (3, 9) we see a pair of numbers, and the order of these is significant; the pair (9, 3) has a different meaning, and in fact is not on the graph. Thus (3, 9) is an ordered pair. The graph consists of a collection of these and is so called 'a set of ordered pairs'. One could thus define the function with graph $y = x^2$ by saying that it is 'the set of ordered pairs of the form (x, x^2)'. This is a rather

pompous and unhelpful way of saying simply that the function is its graph.

This opinion is meant to apply to early instruction in mathematics. At a more advanced level, this definition is significant. For instance, as was mentioned on page 42, functions of a complex variable make us wish we could draw graphs in four dimensions. This definition allows us to define such a graph and to deduce its properties. At this level, and at higher levels, it serves a useful purpose. Teachers will no doubt differ as to how soon they should begin to prepare their pupils' minds for such a definition. The essential point is that pupils should realize that such cumbrous definitions are fabricated for a purpose, and should know at what stage of their mathematics they may expect to receive tangible benefits in return for the additional verbiage.

In Chapter Three we met various examples of mappings. For definiteness, let us imagine plane graph paper, and the transformation T corresponding to reflection in the x-axis. The transformation T sends any vector v to v^*, where $v^* = Tv$. Now T represents a mapping and, so far as we are concerned, *mapping* and *function* are synonyms. So the reflection T is a function. If we were guided by the traditional system of writing a function as $f(x)$, we would have to speak of the reflection $T(v)$. But we do not do this; in Chapter Three, and ever since, we spoke simply of the transformation, or the matrix, T. We keep Tv to denote, not a function, but the point, or vector, to which v is sent by T. Now T is a particularly simple function; it is linear. This, however, does not in any way affect the notation; we could use the same system with any function whatever, and this is very widely done. You may notice in Figure 91 the label on the graph reads 'the function f', and that we write $||f||$ for 'the size of f'. A single letter is used when we are speaking of the whole graph, or of the function specified by that graph. We use the symbol $f(\frac{1}{2})$ to denote a particular number, the height of the graph for $x = \frac{1}{2}$, and $f(x)$ also represents a number, the height of the graph corresponding to x. Note how the familiar $f(x)$ survives in the modern notation; we still speak of the curve with equation $y = f(x)$, which, spelled out in full, means the curve consisting of all the points (x, y) for which $y = f(x)$.

Answers

Chapter Two, page 45

1. F.
2. G and A on first and third paces.
3. D and B on first and second paces.
4. No.
5. Yes, G.
6. $2c-d-e-3f+2g$.
7. E and A on first and second paces.
8. Figure 10 on page 26.

Chapter Three, page 82

1. $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ True.
2. $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}.$
3. Nothing; they are equal.
4. O.
5. $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$
6. Both are $\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}.$
7. $U^2 = I$. True.
8. $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}; \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. No.
9. (1) $\begin{pmatrix} 6 & 4 \\ 4 & 4 \end{pmatrix}$ (2) $\begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}$. Yes, (2) is. Note that (1) is not.
10. Yes. Same answer as question 6.
11. O.
12. O.
13. $E^2-8E = 3I$, so $k = 3$.
14. $F^2+7I = 6F$.
15. $\begin{pmatrix} 7 & 4 \\ 6 & 7 \end{pmatrix}; \begin{pmatrix} 2+k & 4 \\ 6 & 2+k \end{pmatrix}$. Equal if $k = 5$.
16. $q = 3$, $k = 1$.

17. See Chapter Five, final section.
18. Matrix C of question 11.
19. F, question 14.
20. $G = P+2Q+3R+S$; $I = P+S$.
21. Yes.
22. Yes.
23. Yes; 4.
24. It is a linear space of nine dimensions.
25. I.

Chapter Four, page 93

1. (a) $X^* = X,\ Y^* = -Y$.
   (b) $X^* = 2X,\ Y^* = 4Y$.
   (c) $X^* = 2X,\ Y^* = 0$.
2. U; W.

Chapter Five, page 107

The transformation is $a^* = -a,\ b^* = b,\ c^* = -c$, a rotation of 180° about the b-axis. The invariant spaces are the plane $b = 0$ and the line of the b-axis. Every vector in the plane is reversed and has $\lambda = -1$. Every vector in the line is unaltered and has $\lambda = +1$.

Chapter Seven, page 150

1. D^2-1-x^2. No.
2. 1.
3. $X^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, $X^3 = 0$. (1) $\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$ (2) the same. Yes; Yes; polynomials in one matrix behave as in elementary algebra. $I+10X+45X^2$ by the Binomial Theorem. X^3 and all higher powers are 0.
4. $A^2+AB+BA+B^2$; $A^3+AAB+ABA+BAA+ABB+BAB+BBA+B^3$, which equals $A^3+A^2B+ABA+BA^2+AB^2+BAB+B^2A+B^3$.
5. All except (10) and (12). (10) fails, for $2x = 2$ has two solutions, 1 and 6, while $2x = 1$ has none. (12) fails, e.g. $2\times 5 = 0$. A commutative ring with unit element. Yes.
6. Passes all tests, a field. 0 meets the requirements of test (4). The element I for test (9) is the digit 6. Binomial Theorem holds.

7. Sum of two odd numbers is even, so addition is not defined within system; this gets rid of (1) to (5) and (11), (12). (6) to (9) hold; the element I is 1. (10) fails; $5x = 5$ has five solutions, $5x = 1$ none. The system does not fulfil the requirements of any of the types named in this chapter. Binomial Theorem cannot even be stated, since $+$ not defined.

Chapter Nine, page 179

1. It is the rotation with $c = aA+bB$, $s = bA-aB$. This makes $c^2+s^2 = 1$, as required.
2. $X \neq Y$, but $XY = YX = I$. For $XY = M_1M_2M_2M_1 = M_1IM_1 = M_1{}^2 = I$ since $M_2{}^2 = I$ and $M_1{}^2 = I$.
3. A reflection.

Chapter Nine, page 183

1. $V = P+Q+R$.
2. $V = P+2Q+3R$.

published by Penguin Books

Anyone who has read *Mathematician's Delight* or *Prelude to Mathematics* knows W. W. Sawyer as a mathematical lion-tamer. Figures do not merely come to life for him: they eat out of his hand.

In this fourth volume of a series which began with *Vision in Elementary Mathematics*, he discusses those ideas of reform in mathematical education which are very much in the air at present. The slogan 'modern mathematics' is widely heard, although nobody seems to know exactly what it is and some very queer things are being done in its name. Many exponents give no indication of where the 'new mathematics' came from, nor what you can do with it.

A Path to Modern Mathematics shows how various new ideas grow naturally out of the traditional mathematics curriculum and selects those parts of recent mathematical discoveries that look most likely to become of importance to practical men.

Vision in Elementary Mathematics and *The Search for Pattern* (Volumes 1 and 3 of *Introducing Mathematics*) are available in Pelicans. Volume 2 is in preparation.

Cover design by Garth Bell

United Kingdom 30p
Australia $1.00
New Zealand $1.00
South Africa R0.75
Canada $1.25

MATHEMATICS
& STATISTICS
ISBN 0 14
02.0847 X

W. W. Sawyer A Path to Modern Mathematics

ISBN 0 14
02.0847 X